

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 05:55:10 ; Search time 6823 Seconds  
(without alignments)  
6753.761 Million cell updates/sec

Title: US-09-403-882a-1

Sequence: 1 atggccgaggtcagctgcga.....ggcagaagaagcagcttag 951

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:\*

1: gb\_ba:\*

2: gb\_hca:\*

3: gb\_in:\*

4: gb\_lm:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 749.6 | 78.8        | 1256   | 6     | AR175770 Sequence  |
| 2          | 749.6 | 78.8        | 1256   | 6     | AR232005 Sequence  |
| 3          | 749.6 | 78.8        | 1256   | 6     | AX136010 Sequence  |
| 4          | 749.6 | 78.8        | 1256   | 6     | AX136779 Sequence  |
| 5          | 749.6 | 78.8        | 1256   | 6     | AX137830 Sequence  |
| 6          | 749.6 | 78.8        | 1256   | 6     | BD000153 Process F |
| 7          | 749.6 | 78.8        | 1256   | 6     | BD010885 Process F |
| 8          | 737   | 77.5        | 843    | 6     | AX781447 Sequence  |
| 9          | 737   | 77.5        | 843    | 6     | AX795414 Sequence  |
| 10         | 699.4 | 73.5        | 726    | 6     | AX353652 Sequence  |
| 11         | 655.4 | 68.5        | 738    | 6     | A41674 Sequence    |
| 12         | 651   | 68.5        | 711    | 6     | A25748 Variable re |
| 13         | 651   | 68.5        | 711    | 6     | AR051416 Sequence  |
| 14         | 580.6 | 61.1        | 672    | 6     | A25746 Variable re |
| 15         | 580.6 | 61.1        | 672    | 6     | AR051415 Sequence  |
| 16         | 548   | 57.6        | 1725   | 6     | AX001509 Sequence  |
| 17         | 543.6 | 57.2        | 793    | 6     | A41676 Sequence    |
| 18         | 539   | 56.7        | 1668   | 6     | AX001511 Sequence  |
| 19         | 532.4 | 56.0        | 714    | 12    | AY260966 Synthetic |

|    |       |      |      |    |           |                     |
|----|-------|------|------|----|-----------|---------------------|
| 20 | 529   | 55.6 | 720  | 10 | MMU040582 | U40582 Mus musculus |
| 21 | 504.6 | 53.1 | 741  | 12 | AY049714  | AY049714 Synthetic  |
| 22 | 498.2 | 52.4 | 722  | 10 | AF348616  | AF348616 Mus muscu  |
| 23 | 474.4 | 49.9 | 774  | 6  | AX001517  | AX001517 Sequence   |
| 24 | 468.6 | 49.3 | 807  | 6  | BD243923  | BD243923 Gastroint  |
| 25 | 467   | 49.1 | 927  | 6  | AX111695  | AX111695 Sequence   |
| 26 | 465   | 48.9 | 888  | 6  | AX100192  | AX100192 Sequence   |
| 27 | 462.8 | 48.7 | 708  | 6  | A25744    | A25744 Variable re  |
| 28 | 462.8 | 48.6 | 708  | 6  | AR051414  | AR051414 Sequence   |
| 29 | 462.2 | 48.6 | 858  | 6  | AR121618  | AR121618 Sequence   |
| 30 | 462.2 | 48.6 | 870  | 6  | A41893    | A41893 Sequence 1   |
| 31 | 460.6 | 48.4 | 807  | 6  | BD243924  | BD243924 Gastroint  |
| 32 | 460.2 | 48.4 | 807  | 6  | BD243870  | BD243870 Continou   |
| 33 | 460.2 | 48.4 | 807  | 6  | AR317194  | AR317194 Sequence   |
| 34 | 454.8 | 47.8 | 702  | 6  | A42000    | A42000 Sequence 41  |
| 35 | 452   | 47.5 | 756  | 6  | AX954894  | AX954894 Sequence   |
| 36 | 449.6 | 47.3 | 1679 | 6  | AR054409  | AR054409 Sequence   |
| 37 | 449.6 | 47.3 | 1679 | 6  | AR162946  | AR162946 Sequence   |
| 38 | 449.6 | 47.3 | 1679 | 6  | AR202680  | AR202680 Sequence   |
| 39 | 449.6 | 47.3 | 1679 | 6  | AR370694  | AR370694 Sequence   |
| 40 | 449.6 | 47.3 | 1679 | 6  | AX080643  | AX080643 Sequence   |
| 41 | 446   | 46.9 | 735  | 12 | AY725471  | AY725471 Synthetic  |
| 42 | 446   | 46.9 | 759  | 6  | AX001525  | AX001525 Sequence   |
| 43 | 445.4 | 46.8 | 810  | 6  | CQ873224  | CQ873224 Sequence   |
| 44 | 445.4 | 46.8 | 810  | 6  | CQ876881  | CQ876881 Sequence   |
| 45 | 444.2 | 46.7 | 717  | 6  | BD139677  | BD139677 A novel m  |

## ALIGNMENTS

RESULT 1  
LOCUS AR175770 1256 bp. DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 7 from patent US 6309861.  
ACCESSION AR175770  
VERSION AR175770.1 GI:17917069  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1256)  
Ambrosius,D., Rudolph,R., Schaeffner,J. and Schwarz,E.  
PROCESS for the production of naturally folded and secreted  
proteins  
TITILE  
JOURNAL Patent: US 6309861-A 7 30-OCT-2001;  
FEATURES  
source location/Qualifiers  
1..1256  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
Best Local Similarity 98.2%; Pred. No. 3e-189;  
Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 1 ATGGCCGAGGTCAAGCTGCAAGTCAAGGGGAGGCTTAGTGCAGCGTCCCGG 60  
Db 199 ATGGCCGAGGTCAAGCTGCAAGTCTGGGGAGGCTTAGTGCAGCGTCCCGG 258  
QY 61 AAATCTCTCTGTCAGACCTCTGATTCACTTTCAGTAGCTTTGGAATGCACTGGTTCT 120  
Db 259 AAATCTCTCTGTCAGACCTCTGATTCACTTTCAGTAGCTTTGGAATGCACTGGTTCT 318  
QY 121 CAGGCTTCAGAGAGGGCTGAGTGGTTCATATATTATTTAGTGGCACTAGTACATC 180  
Db 319 CAGGCTTCAGAGAGGGCTGAGTGGTTCATATATTATTTAGTGGCACTAGTACATC 378  
QY 181 TACTATGACAGACAGAGGAGGAGGAGTTCACCATCTCCAGAGCAATCCAGAACACC 240  
Db 379 TACTATGACAGACAGAGGAGGAGGAGTTCACCATCTCCAGAGCAATCCAGAACACC 438  
QY 241 CTGTTCTCTGCAATGACCACTTAAAGTCTGAGAGACGGTCAATGATTAATGAGCAAGA 300

Db 439 CTGTTCTCGCAATGATCCAGTCTTAAGGTCTGAGGACACGGCCATGATTATTCGCCGAA 498  
Qy 301 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCGTCTCTCAGAGTGAAGCGGC 360  
Db 499 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCGTCTCTCAGAGTGAAGCGGC 558  
Qy 361 TCAGCGGAGGTGGCTCTGCGGGTGGCGGATCGGACATTGAGCTCAACCGAGTCTCCAGCA 420  
Db 559 TCAGCGGAGGTGGCTCTGCGGGTGGCGGATCGGACATTGAGCTCAACCGAGTCTCCAGCA 618  
Qy 421 ATCATGTGTCATCTCCAGGGGAGAGGGTCAACATGACCTGAGGCCGAGTCAAGTGA 480  
Db 619 ATCATGTCTGCAATCTCCAGGGGAGAGGGTCAACATGACCTGAGGCCGAGTCAAGTGA 678  
Qy 481 AGGTACATGAACTGGTTCCAGAGAGTCAAGGACACTCCGCCAAAAGATGATTTATGAC 540  
Db 679 AGGTACATGAACTGGTTCCAGAGAGTCAAGGACACTCCGCCAAAAGATGATTTATGAC 738  
Qy 541 ACATCCAAACTGTCTTCTGGAAGTCCCTGCTCGCTTCACTGAGGAGTGGAGTCTGGACCTCT 600  
Db 729 ACATCCAAACTGTCTTCTGGAAGTCCCTGCTCGCTTCACTGAGGAGTGGAGTCTGGACCTCT 798  
Qy 601 TACTCTTCACAATAGAGAGATGAGAGGCTGAAGATGCTGCACTTACTACTGCGCAGAG 660  
Db 799 TACTCTTCACAATAGAGAGATGAGAGGCTGAAGATGCTGCACTTACTACTGCGCAGAG 858  
Qy 661 TGGAGTAGTAACCCACTACGTTCCGTTGCTGGAGCCAAAGCTGGAAGACGGGCGGCC 720  
Db 859 TGGAGTAGTAATTCACACTCTTCCGTTGCTGGAGCCAAAGCTGGAAGACGGGCGGCC 918  
Qy 721 GCAGAACAAAATCTCATCTCAGAGAGAGATCTGAATGGGGCGCTGCAGCAAC 772  
Db 919 GCAGAACAAAATCTCATCTCAGAGAGAGATCTGAATGGGGCGCTGCATAGTAAC 970

RESULT 2  
AR232005 1256 bp DNA linear PAT 20-DEC-2002  
LOCUS AR232005  
DEFINITION Sequence 7 from patent US 6455279.  
ACCESSION AR232005  
VERSION AR232005.1 GI:27273847  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1256)  
AUTHORS Ambrosius D., Rudolph R., Schaeffner J. and Schwarz E.  
TITLE Process for the production of naturally folded and secreted proteins by co-secretion of molecular chaperones  
JOURNAL Patent: US 6455279-A 7 24-SEP-2002;  
FEATURES  
source 1..1256  
/organism="Unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
Best Local Similarity 98.2%; Pred. No. 3e-189;  
Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 121 TACTATGCAGACACGTGAAGGGAGGATTACCATCTCCAGAGCAATCCCAAGAACACC 240  
Db 379 TACTATGCAGACACGTGAAGGGAGGATTACCATCTCCAGAGCAATCCCAAGAACACC 438  
Qy 241 CTGTTCTCGCAATGATCCAGTCTTAAGGTCTGAGGACCAAGGTCAACCGTCTCTCAGAGTGAAGCGGC 300  
Db 499 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCGTCTCTCAGAGTGAAGCGGC 498  
Qy 301 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCGTCTCTCAGAGTGAAGCGGC 360  
Db 499 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCGTCTCTCAGAGTGAAGCGGC 558  
Qy 361 TCAGCGGAGGTGGCTCTGCGGGTGGCGGATCGGACATTGAGCTCAACCGAGTCTCCAGCA 420  
Db 559 TCAGCGGAGGTGGCTCTGCGGGTGGCGGATCGGACATTGAGCTCAACCGAGTCTCCAGCA 618  
Qy 421 ATCATGTGTCATCTCCAGGGGAGAGGGTCAACATGACCTGAGGCCGAGTCAAGTGA 480  
Db 619 ATCATGTCTGCAATCTCCAGGGGAGAGGGTCAACATGACCTGAGGCCGAGTCAAGTGA 678  
Qy 481 AGGTACATGAACTGGTTCCAGAGAGTCAAGGACACTCCGCCAAAAGATGATTTATGAC 540  
Db 679 AGGTACATGAACTGGTTCCAGAGAGTCAAGGACACTCCGCCAAAAGATGATTTATGAC 738  
Qy 541 ACATCCAAACTGTCTTCTGGAAGTCCCTGCTCGCTTCACTGAGGAGTGGAGTCTGGACCTCT 600  
Db 729 ACATCCAAACTGTCTTCTGGAAGTCCCTGCTCGCTTCACTGAGGAGTGGAGTCTGGACCTCT 798  
Qy 601 TACTCTTCACAATAGAGAGATGAGAGGCTGAAGATGCTGCACTTACTACTGCGCAGAG 660  
Db 799 TACTCTTCACAATAGAGAGATGAGAGGCTGAAGATGCTGCACTTACTACTGCGCAGAG 858  
Qy 661 TGGAGTAGTAACCCACTACGTTCCGTTGCTGGAGCCAAAGCTGGAAGACGGGCGGCC 720  
Db 859 TGGAGTAGTAATTCACACTCTTCCGTTGCTGGAGCCAAAGCTGGAAGACGGGCGGCC 918  
Qy 721 GCAGAACAAAATCTCATCTCAGAGAGAGATCTGAATGGGGCGCTGCAGCAAC 772  
Db 919 GCAGAACAAAATCTCATCTCAGAGAGAGATCTGAATGGGGCGCTGCATAGTAAC 970

RESULT 3  
AX136010 1256 bp DNA linear PAT 30-MAY-2001  
LOCUS AX136010  
DEFINITION Sequence 7 from Patent EP1054063.  
ACCESSION AX136010  
VERSION AX136010.1 GI:14272438  
KEYWORDS  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
REFERENCE 1  
AUTHORS Ambrosius D., Rudolph R., Schaeffner J. and Schwarz E.  
TITLE Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.  
JOURNAL Patent: EP 1054063-A 7 22-NOV-2000;  
FEATURES  
source F. HOFFMANN-LA ROCHE AG (CH)  
Location/Qualifiers  
1..1256  
/organism="Escherichia coli"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:562"  
199..966  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC39675.1"  
/db\_xref="GI:14272439"  
/translation="MAEVKLOESGGGLVOPGSRKLSAASGFTFSFGMHVROAPE  
KGLRMVAVISGSSFTIYADIVKGRFTISRPNTPLQMTSLRSEDTAMVYCARDY  
GAYVGGCTTYVSSGGGGSGGGGSDIELTOSPIIMASPREKTYTMCSSASV  
RIMNMFQOKSGTSPRMIYDTSKLSGVPARFSGSGSTYSVLTIISMEADAIYYC



## ORIGIN

OQWSSNPLTFGAGTKLELRMAAEOKLISEEDLNGAA"

Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
 Best Local Similarity 98.2%; Pred. No. 3e-189; Indels 0; Gaps 0;  
 Matches 758; Conservative 0; Mismatches 14;

1 ATGGCCGAGGTCAAGCTGAGAGTCAAGGAGGAGCTTATGTCAGGCTGAGAGGTCCTGG 60  
 199 ATGGCCGAGGTCAAGCTGAGAGTCTGGGAGGAGCTTATGTCAGGCTGAGAGGTCCTGG 258  
 61 AAATCTCTCTGTCAGCTCTGATTCATTCTGATGCTTTGGAATGCACTGGTTCTG 120  
 259 AAATCTCTCTGTCAGCTCTGATTCATTCTGATGCTTTGGAATGCACTGGTTCTG 318  
 121 CAGGCTCCAGAGAGAGGAGGCTGAGAGGTCATATATGTCAGGCTGAGAGTACATC 180  
 319 CAGGCTCCAGAGAGAGGAGGCTGAGAGGTCATATATGTCAGGCTGAGAGTACATC 378  
 181 TACTATGACAGACAGTGAAGGAGCAGTTCACCATCTCCAGAGACATCCCAAGAACCC 240  
 379 TACTATGACAGACAGTGAAGGAGCAGTTCACCATCTCCAGAGACATCCCAAGAACCC 438  
 241 CTGTTCTCTCAATATGACAGTCTAAGGTCAGAGACAGGTCATGTAATTAAGTCAAGA 300  
 439 CTGTTCTCTCAATATGACAGTCTAAGGTCAGAGACAGGTCATGTAATTAAGTCAAGA 498  
 301 GATTACGGGGCTTATTTGGGGCCCAAGGACACCGTCTCCAGAGTGAAGGCGGC 360  
 499 GATTACGGGGCTTATTTGGGGCCCAAGGACACCGTCTCCAGAGTGAAGGCGGC 558  
 361 TCAGGCGAGAGTGGCTCTGGCGGTGGCGGATCGGATCGGATTCAGGCTCAAGCTCCAGA 420  
 559 TCAGGCGAGAGTGGCTCTGGCGGTGGCGGATCGGATCGGATTCAGGCTCAAGCTCCAGA 618  
 421 ATCATGTCGATCTTCAGGGAGAGAGGTCACCATGACCTGAGTCCAGTTCAGTGA 480  
 619 ATCATGTCGATCTTCAGGGAGAGAGGTCACCATGACCTGAGTCCAGTTCAGTGA 678  
 481 AGGTACATGAATCTGGTTCACAGAGATGAGGACCTCCCGCCCAAGATGGAATTTATGAC 540  
 679 AGGTACATGAATCTGGTTCACAGAGATGAGGACCTCCCGCCCAAGATGGAATTTATGAC 738  
 541 ACATCCAACTGCTCTGGAAGTCCCTGCTCGCTTCAGTGGAGTGGAGTCTGGAACCTCT 600  
 739 ACATCCAACTGCTCTGGAAGTCCCTGCTCGCTTCAGTGGAGTGGAGTCTGGAACCTCT 798  
 601 TACTCTCTCAATATGACAGATGAGGTCGATGAGATGCTGCACTTAATCTGACAGAG 660  
 799 TACTCTCTCAATATGACAGATGAGGTCGATGAGATGCTGCACTTAATCTGACAGAG 858  
 661 TGGAGTAGTAATCCCACTACAGTTCGCTGGGACCAAGTGAAGTGAACCGGGCGGC 720  
 859 TGGAGTAGTAATCCCACTACAGTTCGCTGGGACCAAGTGAAGTGAACCGGGCGGC 918  
 721 GCAGAACAAAACTCACTCAGAAAGGATCTGAATGGGGCGTCGACGAAAC 772  
 919 GCAGAACAAAACTCACTCAGAAAGGATCTGAATGGGGCGTCGACGAAAC 970

RESULT 4  
 AX136779 1256 bp DNA linear PAT 30-MAY-2001  
 LOCUS AX136779  
 DEFINITION Sequence 7 from Patent EP1077263.  
 ACCESSION AX136779  
 VERSION AX136779.1 GI:14273172  
 KEYWORDS  
 SOURCE Escherichia coli  
 ORGANISM Escherichia coli  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 REFERENCE 1  
 AUTHORS

TITLE Process for producing natural folded and secreted proteins by  
 co-secretion of Chapterones  
 JOURNAL Patent: EP 1077263-A 7 21-FEB-2001;  
 F.HOFPMANN-LA ROCHE AG (CH)

FEATURES  
 source Location/Qualifiers  
 1..1256  
 /organism="Escherichia coli"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:562"  
 199..966  
 /note="unnamed protein product"

## CDS

/codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAC39987.1"  
 /db\_xref="GI:14273173"  
 /translation="MAEVKLQESGGGLVPGGSRKLSCAAGTPTSSFGMHVRYAPR  
 KLEWVAVYSSSSRTVYADYKGRPTSSRDNPXKTLPIQMTSISEPDIAVYCADY  
 GAYWGGTIVYSGGGSGSGGSGSGGSDIELTQSPALMSAPRKRTMTCSASSV  
 RYNNMFQQSGTSPKRWIYDTSLSSGVPAFSGSGSYSLTLTSSMEADNATYYC  
 OQWSSNPLTFGAGTKLELRMAAEOKLISEEDLNGAA"

## ORIGIN

Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
 Best Local Similarity 98.2%; Pred. No. 3e-189; Indels 0; Gaps 0;  
 Matches 758; Conservative 0; Mismatches 14;

1 ATGGCCGAGGTCAAGCTGAGAGTCAAGGAGGAGCTTATGTCAGGCTGAGAGGTCCTGG 60  
 199 ATGGCCGAGGTCAAGCTGAGAGTCTGGGAGGAGCTTATGTCAGGCTGAGAGGTCCTGG 258  
 61 AAATCTCTCTGTCAGCTCTGATTCATTCTGATGCTTTGGAATGCACTGGTTCTG 120  
 259 AAATCTCTCTGTCAGCTCTGATTCATTCTGATGCTTTGGAATGCACTGGTTCTG 318  
 121 CAGGCTCCAGAGAGAGGAGGCTGAGAGGTCATATATGTCAGGCTGAGAGTACATC 180  
 319 CAGGCTCCAGAGAGAGGAGGCTGAGAGGTCATATATGTCAGGCTGAGAGTACATC 378  
 181 TACTATGACAGACAGTGAAGGAGCAGTTCACCATCTCCAGAGACATCCCAAGAACCC 240  
 379 TACTATGACAGACAGTGAAGGAGCAGTTCACCATCTCCAGAGACATCCCAAGAACCC 438  
 241 CTGTTCTCTCAATATGACAGTCTAAGGTCAGAGACAGGTCATGTAATTAAGTCAAGA 300  
 439 CTGTTCTCTCAATATGACAGTCTAAGGTCAGAGACAGGTCATGTAATTAAGTCAAGA 498  
 301 GATTACGGGGCTTATTTGGGGCCCAAGGACACCGTCTCCAGTTCAGTGAAGGCGGC 360  
 499 GATTACGGGGCTTATTTGGGGCCCAAGGACACCGTCTCCAGTTCAGTGAAGGCGGC 558  
 361 TCAGGCGAGAGTGGCTCTGGCGGTGGCGGATCGGATCGGATTCAGGCTCAAGCTCCAGA 420  
 559 TCAGGCGAGAGTGGCTCTGGCGGTGGCGGATCGGATCGGATTCAGGCTCAAGCTCCAGA 618  
 421 ATCATGTCGATCTTCAGGGAGAGAGGTCACCATGACCTGAGTCCAGTTCAGTGA 480  
 619 ATCATGTCGATCTTCAGGGAGAGAGGTCACCATGACCTGAGTCCAGTTCAGTGA 678  
 481 AGGTACATGAATCTGGTTCACAGAGATGAGGACCTCCCGCCCAAGATGGAATTTATGAC 540  
 679 AGGTACATGAATCTGGTTCACAGAGATGAGGACCTCCCGCCCAAGATGGAATTTATGAC 728  
 541 ACATCCAACTGCTCTGGAAGTCCCTGCTCGCTTCAGTGGAGTGGAGTCTGGAACCTCT 600  
 739 ACATCCAACTGCTCTGGAAGTCCCTGCTCGCTTCAGTGGAGTGGAGTCTGGAACCTCT 798  
 601 TACTCTCTCAATATGACAGATGAGGTCGATGAGATGCTGCACTTAATCTGACAGAG 660  
 799 TACTCTCTCAATATGACAGATGAGGTCGATGAGATGCTGCACTTAATCTGACAGAG 858  
 661 TGGAGTAGTAATCCCACTACAGTTCGCTGGGACCAAGTGAAGTGAACCGGGCGGC 720  
 859 TGGAGTAGTAATCCCACTACAGTTCGCTGGGACCAAGTGAAGTGAACCGGGCGGC 918

QY 721 GCAGAACAAAACTCATCTCAGAGAGATGTGAATGGGGCCGTCGACGAC 772  
DB 919 GCAGAACAAAACTCATCTCAGAGAGATGTGAATGGGGCCGTCATGTAAAC 970

RESULT 5  
AX137830  
LOCUS AX137830 1256 bp DNA linear PAT 30-MAY-2001  
DEFINITION Sequence 7 from Patent EP1077262.  
ACCESSION AX137830  
VERSION AX137830.1 GI:14273986  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
F. HOFMANN-LA ROCHE AG (CH)  
FEATURES  
source  
1..1256  
/organism="Escherichia coli"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:562"  
199..966  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC39937.1"  
/db\_xref="GI:14273987"  
/translation="MAEVYLQESGGGLVOPGGSRKLSAASGFTPSFGMHVROAPE  
KLEWVAYISGGSTIYADTKGRFTLSRDPKXTLFLQWMSLRSEPTAYVCARDY  
GAYWQGTIVYSSGGSGGGSGGSDIELTSPALMSRPERKTTMCSASSV  
RYMNFQKSGTSPKRWITYSKLSGGPARSGSGSTYSLTITSSMEADPAITYC  
QOWSNPLTFGAGTLELRKAAAEQKLISEEDLNGAA"

Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
Best Local Similarity 98.2%; Pred. No. 3e-189;  
Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGCCGAGGTCAAGCTGACAGAGTCAAGGGGAGGCTTAGTACAGCTGAGAGTCCCG 60  
DB 199 ATGCCGAGGTCAAGCTGACAGAGTCTGGGGAGGCTTAGTACAGCTGAGAGTCCCG 258

QY 61 AAATCTCTCTGTGACGCTCTGGAATTCAGTTACAGCTTTGGAATGCACTGGGTTGCT 120  
DB 259 AAATCTCTCTGTGACGCTCTGGAATTCAGTTACAGCTTTGGAATGCACTGGGTTGCT 318

QY 121 CAGGCTCAGAGAGAGGGCTGGAGTGGGTCGATATATTAGTAGTGGGAGTACCATC 180  
DB 319 CAGGCTCAGAGAGAGGGCTGGAGTGGGTCGATATATTAGTAGTGGGAGTACCATC 378

QY 181 TACTATGACAGACAGTGAAGGAGCATTCACATCTCCAGAGACATCCCAAGAACACC 240  
DB 379 TACTATGACAGACAGTGAAGGAGCATTCACATCTCCAGAGACATCCCAAGAACACC 438

QY 241 CTGTTCTCTGCAAAATGACAGTCTTAAGTCTGAGACACGGTATATTAATCTGTGCAAGA 300  
DB 439 CTGTTCTCTGCAAAATGACAGTCTTAAGTCTGAGACACGGTATATTAATCTGTGCAAGA 498

QY 301 GATTACGGGGCTTATTGGGGCCAGAGGACACGGTCTCTCCAGGTGGAAGGGCGC 360  
DB 499 GATTACGGGGCTTATTGGGGCCAGAGGACACGGTCTCTCCAGGTGGAAGGGCGT 558

QY 361 TCAGGCGAGAGTGGCTCTGCGGTGGCGGATGCAATGAGCTCAACCAAGTCTCCAGCA 420  
DB 559 TCAGGCGAGAGTGGCTCTGCGGTGGCGGATGCAATGAGCTCAACCAAGTCTCCAGCA 618

QY 421 ATCATGTCTGCATCTCCAGGGAGAGGGTCACTACATGACCTTGACAGTTCAAGTGA 480  
DB 619 ATCATGTCTGCATCTCCAGGGAGAGGGTCACTACATGACCTTGACAGTTCAAGTGA 678

QY 481 AGGTACATGAACCTGGTTCCACAGAACTCAGAGCACCTCCCAAAAGATGATTATGAC 540  
DB 679 AGGTACATGAACCTGGTTCCACAGAACTCAGAGCACCTCCCAAAAGATGATTATGAC 738

QY 541 ACATCAAACTGTCTTCTGAGAGTCCCTGCTGCTTCACTGAGTGGAGTGGGACCTCT 600  
DB 739 ACATCAAACTGTCTTCTGAGAGTCCCTGCTGCTTCACTGAGTGGAGTGGGACCTCT 798

QY 601 TACTCTTCACAAATGACAGATGAGAGCTGAAGTCTGCCACTTACTACTGCCACGAC 660  
DB 799 TACTCTTCACAAATGACAGATGAGAGCTGAAGTCTGCCACTTACTACTGCCACGAC 858

QY 661 TGGAGTAGTAACCCACTACGTTGGGTCTGGGACCAAGTGGAGCTGGAACGGGCGGC 720  
DB 859 TGGAGTAGTAATCCACTACTCTTGGGTCTGGGACCAAGTGGAGCTGGAACGGGCGGC 918

QY 721 GCAGAACAAAACTCATCTCAGAGAGATGTGAATGGGGCCGTCGACGAC 772  
DB 919 GCAGAACAAAACTCATCTCAGAGAGATGTGAATGGGGCCGTCATGTAAAC 970

RESULT 6  
BD000153 1256 bp DNA linear PAT 31-JAN-2002  
LOCUS BD000153  
DEFINITION Process for producing protein that is spontaneously folded and secreted.  
ACCESSION BD000153.1 GI:18623232  
VERSION BD000153.1 JP 2000316591-A/4.  
KEYWORDS unclassified  
SOURCE unclassified  
ORGANISM unclassified  
REFERENCE 1 (bases 1 to 1256)  
AUTHORS Ambrosius, D., Ludl, R., Syefuna, J. and Schwarz, E.  
TITLE Process for producing protein that is spontaneously folded and Patent: JP 2000316591-A 4 21-NOV-2000;  
JOURNAL F. HOFMANN LA ROCHE AG  
COMMENT OS E. coli  
PN 21-NOV-2000  
PD 21-NOV-2000  
PF 25-APR-2000 JP 2000123913  
PR 26-APR-1999 EP 99107412.1  
PI DOROTE AMBROSIOUS, RINER LUDL, JERUG SYEFUNA, ELIZABETH SCHWARZ  
PC C12N15/09, C12P21/02, C12N15/00  
CC  
Key Location/Qualifiers  
FT CDS Location/Qualifiers  
source 1..1256  
/organism="unclassified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
Best Local Similarity 98.2%; Pred. No. 3e-189;  
Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGCCGAGGTCAAGCTGACAGAGTCAAGGGGAGGCTTAGTACAGCTGAGAGTCCCG 60  
DB 199 ATGCCGAGGTCAAGCTGACAGAGTCTGGGGAGGCTTAGTACAGCTGAGAGTCCCG 258

QY 61 AAATCTCTCTGTGACGCTCTGGAATTCAGTTACAGCTTTGGAATGCACTGGGTTGCT 120  
DB 259 AAATCTCTCTGTGACGCTCTGGAATTCAGTTACAGCTTTGGAATGCACTGGGTTGCT 318

QY 121 CAGGCTCAGAGAGAGGGCTGGAGTGGGTCGATATATTAGTAGTGGGAGTACCATC 180  
DB 319 CAGGCTCAGAGAGAGGGCTGGAGTGGGTCGATATATTAGTAGTGGGAGTACCATC 378

QY 181 TACTATGACAGACAGTGAAGGAGCATTCACCATCTCCAGAGACAATCCCAAGAACACC 240  
 DB 379 TACTATGACAGACAGTGAAGGAGCATTCACCATCTCCAGAGACAATCCCAAGAACACC 438  
 QY 241 CTGTTCTCGCAATGACAGCTTAAGGTCTGAGAGACCGGTCAATGATTAATCTGCAAGA 300  
 DB 439 CTGTTCTCGCAATGACAGCTTAAGGTCTGAGAGACCGGTCAATGATTAATCTGCAAGA 498  
 QY 301 GATTACGGGGCTTATTTGGGGCCCAAGGACCAAGGTCAAGCTCTCCAGTGAAGCGCGC 360  
 DB 499 GATTACGGGGCTTATTTGGGGCCCAAGGACCAAGGTCAAGCTCTCCAGTGAAGCGCGC 558  
 QY 361 TCAGCGGAGAGTGGCTCTGGCGGTGGCGAGTCGGAATTTAGAGCTCAACCAAGTCTCAGCA 420  
 DB 559 TCAGCGGAGAGTGGCTCTGGCGGTGGCGAGTCGGAATTTAGAGCTCAACCAAGTCTCAGCA 618  
 QY 421 ATCATGTCTGCATCTCCAGAGGAGAGAGGTCAACATGACCTGCAAGTCCAGTTCAAGTGA 480  
 DB 619 ATCATGTCTGCATCTCCAGAGGAGAGAGGTCAACATGACCTGCAAGTCCAGTTCAAGTGA 678  
 QY 481 AGGTACATGAACCTGTTCCAAAGAGTCAAGGACCTCCCGCAAAAGATGGATTTATGAC 540  
 DB 679 AGGTACATGAACCTGTTCCAAAGAGTCAAGGACCTCCCGCAAAAGATGGATTTATGAC 738  
 QY 541 ACATCCAAACTGCTCTTGGAGTCCCTGCTCGCTTCAATGAGTGGAGTCTGGAGACTCT 600  
 DB 739 ACATCCAAACTGCTCTTGGAGTCCCTGCTCGCTTCAATGAGTGGAGTCTGGAGACTCT 798  
 QY 601 TACTCTTCACATCAATGACAGATGAGGCTGAAGATGCTGCACTTAATTAATCTCCAGCAG 660  
 DB 799 TACTCTTCACATCAATGACAGATGAGGCTGAAGATGCTGCACTTAATTAATCTCCAGCAG 858  
 QY 661 TGGAGTAGTAAACCAATCAAGTTCGCTGGAGCCAAAGTGGAGTGAAGCGGGGGCC 720  
 DB 859 TGGAGTAGTAAACCAATCAAGTTCGCTGGAGCCAAAGTGGAGTGAAGCGGGGGCC 918  
 QY 721 GCAGAACAAATACTCATCTCAGAGAGATCTGAATGGGGCGCTGCAGCAG 772  
 DB 919 GCAGAACAAATACTCATCTCAGAGAGATCTGAATGGGGCGCTGCAGCAG 970  
 RESULT 7  
 LOCUS BD010885 1256 bp DNA linear PAT 31-JAN-2002  
 DEFINITION Process for producing spontaneously folded secreted protein by  
 co-secretion of molecular chaperone.  
 ACCESSION BD010885  
 VERSION BD010885.1 GI:18639258  
 KEYWORDS JP 2001061487-A/4.  
 SOURCE Escherichia coli  
 ORGANISM Escherichia coli  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Escherichia.  
 1 (bases 1 to 1256)  
 Ambrosius, D., Ludl, R., Schaffner, J. and Schwarz, B.  
 Process for producing spontaneously folded secreted protein by  
 co-secretion of molecular chaperone  
 Patent: JP 2001061487-A 4 13-MAR-2001;  
 F. HOFMANN LA ROCHE AG  
 OS Escherichia coli  
 JOURNAL  
 COMMENT  
 PD 13-MAR-2001  
 PN JP 2001061487-A/4  
 PP 31-JUL-2000 JP 2000231804  
 PR 29-JUL-1999 EP 9911481.5  
 PI DOROCY AMBROSIOUS, RINER LUDL, JAEK SCHAFFNER, ELIZABETH SCHWARZ  
 PC C12N15/09, C12N9/48, C12P21/02, C12P21/08, C12N1/21, PC  
 (C12N9/48, C12R1:19), (C12P21/02, C12R1:19), (C12P21/08, C12R1:19), PC  
 C12N15/00  
 CC  
 FH Key Location/Qualifiers  
 FT CDS (199)..(969).  
 Location/Qualifiers

source 1..1256  
 /organism="Escherichia coli"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:562"  
 ORIGIN  
 Query Match 78.8%; Score 749.6; DB 6; Length 1256;  
 Best Local Similarity 98.2%; Pred. No. 3e-189;  
 Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 ATGGCCGAGGTCAAGTGCAGAGTCAAGGAGGCTTATGATGACCTGAGAGGCTCCCG 60  
 DB 199 ATGGCCGAGGTCAAGTGCAGAGTCAAGGAGGCTTATGATGACCTGAGAGGCTCCCG 258  
 QY 61 AAATCTCCGTCGACGCTCTGGAATTCACCTTCAGTACCTTGGAAATGACCTGGATTCGT 120  
 DB 259 AAATCTCCGTCGACGCTCTGGAATTCACCTTCAGTACCTTGGAAATGACCTGGATTCGT 318  
 QY 121 CAGGCTCAGAGAGGGGCTGAGTGGGTCCGATATATTAATGAGTGAAGTATACATC 180  
 DB 319 CAGGCTCAGAGAGGGGCTGAGTGGGTCCGATATATTAATGAGTGAAGTATACATC 378  
 QY 181 TACTATGACAGACAGTGAAGGAGCATTCACCATCTCCAGAGACAATCCCAAGAACACC 240  
 DB 379 TACTATGACAGACAGTGAAGGAGCATTCACCATCTCCAGAGACAATCCCAAGAACACC 438  
 QY 241 CTGTTCTCGCAATGACAGCTTAAGGTCTGAGAGACCGGTCAATGATTAATCTGCAAGA 300  
 DB 439 CTGTTCTCGCAATGACAGCTTAAGGTCTGAGAGACCGGTCAATGATTAATCTGCAAGA 498  
 QY 301 GATTACGGGGCTTATTTGGGGCCCAAGGACCAAGGTCAAGCTCTCCAGTGAAGCGCGC 360  
 DB 499 GATTACGGGGCTTATTTGGGGCCCAAGGACCAAGGTCAAGCTCTCCAGTGAAGCGCGC 558  
 QY 361 TCAGCGGAGAGTGGCTCTGGCGGTGGCGAGTCGGAATTTAGAGCTCAACCAAGTCTCAGCA 420  
 DB 559 TCAGCGGAGAGTGGCTCTGGCGGTGGCGAGTCGGAATTTAGAGCTCAACCAAGTCTCAGCA 618  
 QY 421 ATCATGTCTGCATCTCCAGAGGAGAGAGGTCAACATGACCTGCAAGTCCAGTTCAAGTGA 480  
 DB 619 ATCATGTCTGCATCTCCAGAGGAGAGAGGTCAACATGACCTGCAAGTCCAGTTCAAGTGA 678  
 QY 481 AGGTACATGAACCTGTTCCAAAGAGTCAAGGACCTCCCGCAAAAGATGGATTTATGAC 540  
 DB 679 AGGTACATGAACCTGTTCCAAAGAGTCAAGGACCTCCCGCAAAAGATGGATTTATGAC 738  
 QY 541 ACATCCAAACTGCTCTTGGAGTCCCTGCTCGCTTCAATGAGTGGAGTCTGGAGACTCT 600  
 DB 739 ACATCCAAACTGCTCTTGGAGTCCCTGCTCGCTTCAATGAGTGGAGTCTGGAGACTCT 798  
 QY 601 TACTCTTCACATCAATGACAGATGAGGCTGAAGATGCTGCACTTAATTAATCTCCAGCAG 660  
 DB 799 TACTCTTCACATCAATGACAGATGAGGCTGAAGATGCTGCACTTAATTAATCTCCAGCAG 858  
 QY 661 TGGAGTAGTAAACCAATCAAGTTCGCTGGAGCCAAAGTGGAGTGAAGCGGGGGCC 720  
 DB 859 TGGAGTAGTAAACCAATCAAGTTCGCTGGAGCCAAAGTGGAGTGAAGCGGGGGCC 918  
 QY 721 GCAGAACAAATACTCATCTCAGAGAGATCTGAATGGGGCGCTGCAGCAG 772  
 DB 919 GCAGAACAAATACTCATCTCAGAGAGATCTGAATGGGGCGCTGCAGCAG 970  
 RESULT 8  
 LOCUS AX781447 843 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Sequence 3 from Patent EP1321524.  
 ACCESSION AX781447  
 VERSION AX781447.1 GI:32949302  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Mahn, A., Hanke, S. and Petesch, D.  
TITLE Method of increasing the transgene-coded biomolecule content in  
JOURNAL Organisms  
Patent: EP 1321524-A 3 25-JUN-2003;  
Duering, Klaus, Dr. (DE)  
FEATURES Location/Qualifiers  
source 1..843  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description of the Artificial Sequence:  
scFv-antibody"

ORIGIN  
Query Match 77.5%; Score 737; DB 6; Length 843;  
Best Local Similarity 97.4%; Pred. No. 7.1e-186;  
Matches 749; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

4 GCCGAGTCAAGCTGACAGAGTCAGGGGAGCTTAAAGTCAAGCTGAGAGGTCCTCGGAAA 63  
70 GCTGATGTCAGCTGAGTCTGAGGAGGCTTAAAGTCAAGCTGAGAGGTCCTCGGAAA 129  
64 CTCTCTGTCAGCTCTGATTTCACTTCAAGTGAATGCACTGGGTTGTCAG 123  
130 CTCTCTGTCAGCTCTGATTTCACTTCAAGTGAATGCACTGGGTTGTCAG 189  
124 GCTCAGAGAGGGGCTGAGTGGGTGCAATATTAGTAGGAGTAGTCAATCTAC 183  
190 GCTCAGAGAGGGGCTGAGTGGGTGCAATATTAGTAGGAGTAGTCAATCTAC 249  
184 TATGACAGACAGTGAAGGAGATTCACATCTCCAGAGCAATCCCAAGAACCTCTG 243  
250 TATGACAGACAGTGAAGGAGATTCACATCTCCAGAGCAATCCCAAGAACCTCTG 309  
244 TTCTGCAAAATGACAGTCTAAGTCTGAGAGCAAGGTCATGATTTACTGCAAGAGAT 303  
310 TTCTGCAAAATGACAGTCTAAGTCTGAGAGCAAGGTCATGATTTACTGCAAGAGAT 369  
304 TACGGGGCTTATTGGGGCCAAAGGACCAAGTCAAGCTCTCTCAGTGGAGGCGGCTCA 363  
370 TACGGGGCTTATTGGGGCCAAAGGACCAAGTCAAGCTCTCTCAGTGGAGGCGGCTCA 429  
364 GCGGAGGTGGCTCTGGGCGGTGGAGATTCAGTCAAGTCAAGTCTCTCAGTCAATC 423  
430 GCGGAGGTGGCTCTGGGCGGTGGAGATTCAGTCAAGTCAAGTCTCTCAGTCAATC 489  
424 ATGTCTGATCTCCAGGGGAGAGGTCACATGACCTGAGTGCAGTTCAAGTGAAG 483  
490 ATGTCTGATCTCCAGGGGAGAGGTCACATGACCTGAGTGCAGTTCAAGTGAAG 549  
484 TACATGAACCTGGTCCAAAGAGTCAAGGACCTCTCCCAAAAGATGATTTATGACACA 543  
550 TACATGAACCTGGTCCAAAGAGTCAAGGACCTCTCCCAAAAGATGATTTATGACACA 609  
544 TCCAAACTGTCTTTCGAGATCCCTGCTCCTTCAGTGGAGGAGTGGAGCTCTTTAC 603  
610 TCCAAACTGTCTTTCGAGATCCCTGCTCCTTCAGTGGAGGAGTGGAGCTCTTTAC 669  
604 TCTCTCAAAATGACAGTGAAGGCTGAAGTGCCTTCACTTCACTGAGCAGAGTGG 663  
670 TCTCTCAAAATGACAGTGAAGGCTGAAGTGCCTTCACTTCACTGAGCAGAGTGG 729  
664 AGTAGTAACCACTCAAGTTCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 723  
730 AGTAGTAACCACTCAAGTTCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 789  
724 GAACAAAAATCTATCTCAGAAAGAGATTCGATGGGCGGCTGAGGAAC 772  
790 GAACAAAAATCTATCTCAGAAAGAGATTCGATGGGCGGCTGAGGAAC 838

AX795414  
LOCUS AX795414 843 bp DNA linear PAT 04-OCT-2003  
DEFINITION Sequence 3 from Patent WO03052109.  
ACCESSION AX795414  
VERSION AX795414.1 GI:37516087  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Duering, K., Mahn, A., Hanke, S. and Petesch, D.  
TITLE Method of increasing the transgene-coded biomolecule content in  
JOURNAL Organisms  
Patent: WO 03052109-A 3 26-JUN-2003;  
Duering, Klaus (DE)  
FEATURES Location/Qualifiers  
source 1..843  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="scFv-antibody"

ORIGIN  
Query Match 77.5%; Score 737; DB 6; Length 843;  
Best Local Similarity 97.4%; Pred. No. 7.1e-186;  
Matches 749; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

4 GCCGAGTCAAGCTGACAGAGTCAGGGGAGCTTAAAGTCAAGCTGAGAGGTCCTCGGAAA 63  
70 GCTGATGTCAGCTGAGTCTGAGGAGGCTTAAAGTCAAGCTGAGAGGTCCTCGGAAA 129  
64 CTCTCTGTCAGCTCTGATTTCACTTCAAGTGAATGCACTGGGTTGTCAG 123  
130 CTCTCTGTCAGCTCTGATTTCACTTCAAGTGAATGCACTGGGTTGTCAG 189  
124 GCTCAGAGAGGGGCTGAGTGGGTGCAATATTAGTAGGAGTAGTCAATCTAC 183  
190 GCTCAGAGAGGGGCTGAGTGGGTGCAATATTAGTAGGAGTAGTCAATCTAC 249  
184 TATGACAGACAGTGAAGGAGATTCACATCTCCAGAGCAATCCCAAGAACCTCTG 243  
250 TATGACAGACAGTGAAGGAGATTCACATCTCCAGAGCAATCCCAAGAACCTCTG 309  
244 TTCTGCAAAATGACAGTCTAAGTCTGAGAGCAAGGTCATGATTTACTGCAAGAGAT 303  
310 TTCTGCAAAATGACAGTCTAAGTCTGAGAGCAAGGTCATGATTTACTGCAAGAGAT 369  
304 TACGGGGCTTATTGGGGCCAAAGGACCAAGTCAAGCTCTCTCAGTGGAGGCGGCTCA 363  
370 TACGGGGCTTATTGGGGCCAAAGGACCAAGTCAAGCTCTCTCAGTGGAGGCGGCTCA 429  
364 GCGGAGGTGGCTCTGGGCGGTGGAGATTCAGTCAAGTCAAGTCTCTCAGTCAATC 423  
430 GCGGAGGTGGCTCTGGGCGGTGGAGATTCAGTCAAGTCAAGTCTCTCAGTCAATC 489  
424 ATGTCTGATCTCCAGGGGAGAGGTCACATGACCTGAGTGCAGTTCAAGTGAAG 483  
490 ATGTCTGATCTCCAGGGGAGAGGTCACATGACCTGAGTGCAGTTCAAGTGAAG 549  
484 TACATGAACCTGGTCCAAAGAGTCAAGGACCTCTCCCAAAAGATGATTTATGACACA 543  
550 TACATGAACCTGGTCCAAAGAGTCAAGGACCTCTCCCAAAAGATGATTTATGACACA 609  
544 TCCAAACTGTCTTTCGAGATCCCTGCTCCTTCAGTGGAGGAGTGGAGCTCTTTAC 603  
610 TCCAAACTGTCTTTCGAGATCCCTGCTCCTTCAGTGGAGGAGTGGAGCTCTTTAC 669  
604 TCTCTCAAAATGACAGTGAAGGCTGAAGTGCCTTCACTTCACTGAGCAGAGTGG 663  
670 TCTCTCAAAATGACAGTGAAGGCTGAAGTGCCTTCACTTCACTGAGCAGAGTGG 729  
664 AGTAGTAACCACTCAAGTTCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 723

```

Db      730 AGTAGATTCACCTCTTGGTGTCTGGAGCAACAGCTGAGCTGAACGGGCGGCGCA 789
Qy      724 GAACAAAATCATCTCAGAAAGATCTGAATGGGCGCTGACGCAAC 772
Db      790 GAACAAAATCATCTCAGAAAGATCTGAATGGATTCGAAGACGAC 838

RESULT 10
AX353652
LOCUS      AX353652             726 bp      DNA      linear      PAT 06-FEB-2002
DEFINITION Sequence 1 from Patent WO0204020.
ACCESSION  AX353652
VERSION     AX353652.1   GI:18618722
KEYWORDS
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.

REFERENCE   1
AUTHORS    Duerling, K. and Brinkmann, O.
TITLE       Pathogen resistance in organisms
JOURNAL     Patent: WO 0204020-A 1 17-JAN-2002;
            MPB Cologne GmbH (DE)
FEATURES
            Location/Qualifiers
            1..726
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Beschreibung der kuenstlichen Sequenz:
               beFV-Antikuerper, der 2-Phenylloxazol-5-on bindet"

ORIGIN
Query Match      73.5%; Score 699.4; DB 6; Length 726;
Best Local Similarity 97.8%; Pred. No. 8e-176;
Matches 709; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy      1  ATGGCCGAGCTCAAGCTGACAGAGTCAAGGCGGAGCTTACTGAGCTTGAGGGGTCGCCG 60
Db      1  ATGGCTGATGTGACAGCTGGTGTGAGTCTGGGGGAGGCTTATGACGCTGAGGGGTCGCCG 60
Qy      61  AAACCTCTCTGTGACACCTCTGTGATTCACCTTACAGTAGCTTTGGAAATGACATGGGTTCTG 120
Db      61  AAACCTCTCTGTGACACCTCTGTGATTCACCTTACAGTAGCTTTGGAAATGACATGGGTTCTG 120
Qy      121  CAGGCTCCAGAGAAAGGGGCTGAGTGGGTGCATATATTAGTAGTGAGTAGTACCATC 180
Db      121  CAGGCTCCAGAGAAAGGGGCTGAGTGGGTGCATATATTAGTAGTGAGTAGTACCATC 180
Qy      181  TACTATGACAGACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACC 240
Db      181  TACTATGACAGACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACC 240
Qy      241  CTGTCTCTGCAATGACCACTTAAGTCTGAGAGACAGGTCATGTTTACTGTGCAAG 300
Db      241  CTGTCTCTGCAATGACCACTTAAGTCTGAGAGACAGGTCATGTTTACTGTGCAAG 300
Qy      301  GATTACGGGGCTAATGGGGGCAAGGGACCAAGGTACCGCTCCACAGTGAAGGGCGC 360
Db      301  GATTACGGGGCTAATGGGGGCAAGGGACCAAGGTACCGCTCCACAGTGAAGGGCGC 360
Qy      361  TCAGGGGAGAGTGGCTCTGGCGGTGGCGAATCGGACATTGAGAGTCAACCAAGTCTCCAGCA 420
Db      361  TCAGGGGAGAGTGGCTCTGGCGGTGGCGAATCGGACATTGAGAGTCAACCAAGTCTCCAGCA 420
Qy      421  ATCATGTCTGCATCTCCAGGGGAGAGGCTCAACATGACCTGAGTGCAGATTCAAGTGA 480
Db      421  ATCATGTCTGCATCTCCAGGGGAGAGGCTCAACATGACCTGAGTGCAGATTCAAGTGA 480
Qy      481  AGGTACATGAATCGTTCCAAAGAGTCAAGGACCTCCCAAAAGATGATTTATGAC 540
Db      481  AGGTACATGAATCGTTCCAAAGAGTCAAGGACCTCCCAAAAGATGATTTATGAC 540
Qy      541  ACATCCAAATGTTCTTGGAGTCCCTGCTCGCTTCAGTGAGGCAATGGGTCGGGACCTCT 600
Db      541  ACATCCAAATGTTCTTGGAGTCCCTGCTCGCTTCAGTGAGGCAATGGGTCGGGACCTCT 600

```

```

Db      541 ACATCCAAATGTTCTTGGAGTCCCTGCTCGCTTCAGTGAGGCAATGGGTCGGGACCTCT 600
Qy      601  TACTCTCTCACAATGACAGATGAGAGCTGAAGATGTGCCACTTACTACTGCCAGAG 660
Db      601  TACTCTCTCACAATGACAGATGAGAGCTGAAGATGTGCCACTTACTACTGCCAGAG 660
Qy      661  TGGAGTAGTAACCCACTCAAGTTGGTGTGGGACCAAGCTGAGAGTGAACGGGCGGCC 720
Db      661  TGGAGTAGTAATCCACTCAAGTTGGTGTGGGACCAAGCTGAGAGTGAACGGGCGGCC 720
Qy      721  GCAGA 725
Db      721  GCATA 725

RESULT 11
A41674
LOCUS      A41674             738 bp      DNA      linear      PAT 05-MAR-1997
DEFINITION Sequence 1 from Patent EP0630968.
ACCESSION  A41674
VERSION     A41674.1   GI:2297297
KEYWORDS
SOURCE      unidentified
ORGANISM    unidentified.

REFERENCE   1
AUTHORS    Fritz, H., Hennecke, F. and Kolmar, H.
TITLE       Genetic selection of proteins able to bind a ligand by
            signal-transduction in a microorganism
JOURNAL     Patent: EP 0630968-A 1 28-DEC-1994;
            BEHRINGWERKE AG (DE)
COMMENT     Other publication JP 6343472 941220
            Other publication AU 6461494 941215
            Other publication CA 2125536 941211
            Other publication DE 4319296 941215.

FEATURES
            Location/Qualifiers
            1..738
               /organism="unidentified"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32644"
               /note="unnamed protein product; Protein sequence is in
               conflict with the conceptual translation"
               /codon_start=1
               /protein_id="CAA02622.1"
               /db_xref="GI:2297298"
               /translation="QRVDYVQLVEVGSGFVQPGSRKLSAASGTFPSSFGMHWROA
               PERGLEWVAIVISSSTIYADTVKGRFTISDNPNKTLFLQWTSIRSEDTMYTCAR
               DYGAIVWQDGLTVVSGSGSGSGSGSGSGSIVLTQSPALMSASPGKRVMTCSASS
               SYRWMMVFOOKSGTSPKRWYDTSKLSGVPARFSGSGSGTSLTISMEADATY
               YCQWMSNPLFTGAGTKLELKR"

ORIGIN
Query Match      68.9%; Score 655.4; DB 6; Length 738;
Best Local Similarity 95.0%; Pred. No. 4.7e-164;
Matches 677; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy      2  TGGCCAGAGTCAAGCTGACAGAGTCAAGGGGAGGCTTAATGTCAGCCTTGAGAGGTCGCCGA 61
Db      8  TCACATCACTGACAGCTGAGTGAAGCTTGGGGAGGCTTTGTGACGCTTGAGAGGTCGCCGA 67
Qy      62  AACTCTCTGTGAGAGCTCTGATTCACCTTCAAGTAGCTTGAATGCACTGGGTTGCTC 121
Db      62  AACTCTCTGTGAGAGCTCTGATTCACCTTCAAGTAGCTTGAATGCACTGGGTTGCTC 127
Qy      68  AACTCTCTGTGAGAGCTCTGATTCACCTTCAAGTAGCTTGAATGCACTGGGTTGCTC 127
Db      122  AGGCTCCAGAGAAAGGGCTGAGTGGGTGCATATATTAGTAGTGCAGTAGTACATCT 181
Qy      128  AGGCTCCAGAGAAAGGGCTGAGTGGGTGCATATATTAGTAGTGCAGTAGTACATCT 187
Db      182  ACTATGCAACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACC 241
Qy      188  ACTATGCAACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACC 247
Db      188  ACTATGCAACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACC 247

```

QY 242 TGTTCCTGCAATGACCACTTAAGTCTGAGAGACCGGTCACTATTACTGTGCAAG 301  
Db 248 TGTTCCTGAGATGACCACTTAAGTCTGAGAGACCGGTCACTATTACTGTGCAAG 307  
QY 302 ATTACGGGGCTTAATGAGGGGCAAGGGACCAAGGTCAAGGTCTCTCAGGTGAGGGGCT 361  
Db 308 ATTACGGGGCTTAATGAGGGGCAAGGGACCAAGGTCTCTCAGGTGAGGGGCT 367  
QY 362 CAGGCGAGAGTGGCTCTGCGGCTGCGAGATCGGACATTGAGCTCAACCAAGTCTCCAGCAA 421  
Db 368 CAGGCGAGAGTGGCTCTGCGGCTGCGAGATCGGACATTGAGCTCTCAGGTGAGGGGCT 427  
QY 422 TCATGTCTGCAATCTCCAGGGAGAGAGGTACCATGACCTGACAGTGTCAAGTGTAA 481  
Db 428 TCATGTCTGCAATCTCCAGGGAGAGAGGTACCATGACCTGACAGTGTCAAGTGTAA 487  
QY 482 GGTACATGAATGTTTCCAGAGAAAGTCAAGGACCTCCCAAAAGATGATTTATGACA 541  
Db 488 GGTACATGAATGTTTCCAGAGAAAGTCAAGGACCTCCCAAAAGATGATTTATGACA 547  
QY 542 CATCCAACTGTCTTCTGAGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTT 601  
Db 548 CATCCAACTGTCTTCTGAGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTTT 607  
QY 602 ACTCTCTCAATCAAGCAGATGAGAGGTGAGAGTGTCTCCCTTAATCTGCTCCAGCAGT 661  
Db 608 ACTCTCTCAATCAAGCAGATGAGAGGTGAGAGTGTCTCCCTTAATCTGCTCCAGCAGT 667  
QY 662 GGAGTAGTAACCACTCAAGTCTGAGTCTGGAGCCAGGTGAGGTGAGAAACG 714  
Db 668 GGAGTAGTAACCACTCAAGTCTGAGTCTGGAGCCAGGTGAGGTGAGAAACG 720

RESULT 12  
LOCUS A25748 711 bp DNA linear PAT 14-MAR-1995  
DEFINITION Variable region heavy and kappa light chain linked genes (stage 3).  
ACCESSION A25748  
VERSION A25748.1 GI:904741  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 711)  
AUTHORS  
TITLE TREATMENT OF CELL POPULATIONS  
JOURNAL Patent: WO 9303151-A 59 18-FEB-1993;  
FEATURES  
source location/Qualifiers  
1..711  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="NQ10/12.5-stage 2"  
1..345  
V\_region /product="variable region of Ig heavy chain"  
346..387  
misc\_feature /note="linker peptide"  
388..711  
V\_region /product="variable region of Ig kappa light chain"  
ORIGIN

Query Match 68.5%; Score 651; DB 6; Length 711;  
Best Local Similarity 95.0%; Pred. No. 7.1e-163;  
Matches 672; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 7 GAGGTCAAGCTGACAGAGTCAAGGGAGGCTTAAGTCAAGGCTGAGAGGCTCCGGAAACTC 66  
Db 1 GATGTGACGCTGTGAGAGTCTGGGGAGGCTTAAGTCAAGGCTGAGAGGCTCCGGAAACTC 60  
QY 67 TCCTGTGACGCTCTGAGATTCATTTCAAGTACCTTTGGAATGCACTGGGTTGTCAGGCT 126  
Db 61 TCCTGTGACGCTCTGAGATTCATTTCAAGTACCTTTGGAATGCACTGGGTTGTCAGGCT 120

QY 127 CCAGGAAGGGGCTGAGAGTGGGTGSCATATATTAGTAGTGGCAGTAGTACCATCTACTAT 186  
Db 121 CCAGGAAGGGGCTGAGAGTGGGTGSCATATATTAGTAGTGGCAGTAGTACCATCTACTAT 180  
QY 187 GCAGACACAGTGAAGGGGAGCATTCACCATCTCCAGAGACATTCACCAAGAACCCCTGTTT 246  
Db 181 GCAGACACAGTGAAGGGGAGCATTCACCATCTCCAGAGACATTCACCAAGAACCCCTGTTT 240  
QY 247 CTGCAATGACCAAGTCTAAGTCTGAGAGACACGTCATGTATTTACTGTGCAAGATTAAC 306  
Db 241 CTGCAATGACCAAGTCTAAGTCTGAGAGACACGTCATGTATTTACTGTGCAAGATTAAC 300  
QY 307 GGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCTCTCTCAAGTGTGAGGGGCTCAAGC 366  
Db 301 GGGGCTTATTTGGGGCCAAAGGACCAAGGTCAACCTCTCTCAAGTGTGAGGGGCTCAAGC 360  
QY 367 GGAAGTGGCTCTGGGCGGTGGGGAGATCGGACATTTAGTCACTCCAGTCTCCAGCAATCATG 426  
Db 361 GGAAGTGGCTCTGGGCGGTGGGGAGATCGGACATTTAGTCACTCCAGTCTCCAGCAATCATG 420  
QY 427 TCTGATCTCCAGGGGAGAGGATCCATGACCTGACAGTGCAGTTCAAGTGAAGTAC 486  
Db 421 TCTGATCTCCAGGGGAGAGGATCCATGACCTGACAGTGCAGTTCAAGTGAAGTAC 480  
QY 487 ATGAACCTGTTCCACAGAAAGTCAAGGACCTTCCCAAAAGATGATTTATGACATTC 546  
Db 481 ATGAACCTGTTCCACAGAAAGTCAAGGACCTTCCCAAAAGATGATTTATGACATTC 540  
QY 547 AAACGTCTTCTGAGAGTCCCTGCTCGCTTCAGTGGCAGTGGGACCTCTACTCT 606  
Db 541 AAACGTCTTCTGAGAGTCCCTGCTCGCTTCAGTGGCAGTGGGACCTCTACTCT 600  
QY 607 CTCAATCAAGCAGATGAGAGGCTGAAGATCTGCACTTAATCTGCAAGCAGTGAAGT 666  
Db 601 CTCAATCAAGCAGATGAGAGGCTGAAGATCTGCACTTAATCTGCAAGCAGTGAAGT 660  
QY 667 AGTAACCACTCAAGTGTGGTGTGGGACCAAGCTGAGAGCTGAAACG 713  
Db 661 AGTAACCACTCAAGTGTGGTGTGGGACCAAGCTGAGAGCTGAAACG 707

RESULT 13  
LOCUS AR051416 711 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 64 from patent US 5830663.  
ACCESSION AR051416  
VERSION AR051416.1 GI:5974780  
KEYWORDS  
SOURCE unknown.  
ORGANISM unknown.  
REFERENCE 1 (bases 1 to 711)  
AUTHORS Emblemson,M.J., Gorochev,G., Jones,P.T. and Winter,G.P.  
TITLE In situ recombinant PCR within single cells  
JOURNAL Patent: US 5830663-A 64 03-NOV-1998;  
FEATURES  
source location/Qualifiers  
1..711  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN

Query Match 68.5%; Score 651; DB 6; Length 711;  
Best Local Similarity 95.0%; Pred. No. 7.1e-163;  
Matches 672; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 7 GAGGTCAAGCTGACAGAGTCAAGGGAGGCTTAAGTCAAGGCTGAGAGGCTCCGGAAACTC 66  
Db 1 GATGTGACGCTGTGAGAGTCTGGGGAGGCTTAAGTCAAGGCTGAGAGGCTCCGGAAACTC 60  
QY 67 TCCTGTGACGCTCTGAGATTCATTTCAAGTACCTTTGGAATGCACTGGGTTGTCAGGCT 126  
Db 61 TCCTGTGACGCTCTGAGATTCATTTCAAGTACCTTTGGAATGCACTGGGTTGTCAGGCT 120

QY 127 CCAGAGAGAGGGCTGGAGTGGGTGCGATATATTAGTAGTGGCAGTAGTACCATCTATCTAT 186  
DB 121 CCAGAGAGAGGGCTGGAGTGGGTGCGATATATTAGTAGTGGCAGTAGTACCATCTATCTAT 180  
QY 187 GGAGACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACCCGTTC 246  
DB 181 GGAGACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACCCGTTC 240  
QY 247 CTGCAATATGACAGTCTAAGTCTGAGGACACGGTCATGTATTAAGTGCAGAGATTAC 306  
DB 241 CTGCAATATGACAGTCTAAGTCTGAGGACACGGTCATGTATTAAGTGCAGAGATTAC 300  
QY 307 GGGGCTTATTGGGGCCAAAGGAGACACAGTCACTGCTCTCAGTGGAGGCGGCTCAGGC 366  
DB 301 GGGGCTTATTGGGGCCAAAGGAGACACAGTCACTGCTCTCAGTGGAGGCGGCTCAGGC 360  
QY 367 GGAGGTGGCTCTGGCGGCTGGCGGATGGGACATTGAGCTCAACCGTCTCCAGAAATCATG 426  
DB 361 GGGTGGGAGTGGCGGCGGCTCTCAAAATGTTCTCAACCGTCTCCAGAAATCATG 420  
QY 427 TCTGCATCTCCAGGGGAGAGGTCACCATGACCTGGAGTGGCAGTTCAGATGTAAGTAC 486  
DB 421 TCTGCATCTCCAGGGGAGAGGTCACCATGACCTGGAGTGGCAGTTCAGATGTAAGTAC 480  
QY 487 ATGAACGTGTTCCAAAGAGTCAAGGACACCTCCCAAAAGATGATTATGACACATCC 546  
DB 481 ATGAACGTGTTCCAAAGAGTCAAGGACACCTCCCAAAAGATGATTATGACACATCC 540  
QY 547 AAACGTCTTTTGGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGGACCTTTACTCT 606  
DB 541 AAACGTCTTTTGGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGGACCTTTACTCT 600  
QY 607 CTCACAAATGACGAGTGAAGGCTGAAGATGCTGCCACTTACTACTGCCAGAGTGAAGT 666  
DB 601 CTCACAAATGACGAGTGAAGGCTGAAGATGCTGCCACTTACTACTGCCAGAGTGAAGT 660  
QY 667 AGTAACCACTCACTGCTGGTGGTGGAGCAAGCTGAGAGTGAAGC 713  
DB 661 AGTAACCACTCACTGCTGGTGGTGGAGCAAGCTGAGAGTGAAGC 707

RESULT 14  
A25746 672 bp DNA linear PAT 14-MAR-1995  
LOCUS A25746  
DEFINITION Variable region heavy and kappa light chain linked genes (stage 2).  
ACCESSION A25746  
VERSION A25746.1 GI:904740  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 672)  
TREATMENT OF CELL POPULATIONS  
AUTHORS Patent: WO 9303151-A 57 18-FEB-1993;  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..672  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="NO10/12.5-stage 1"  
V\_region 1..342  
/product="variable region of Ig heavy chain"  
misc\_feature 343..348  
/note="linker peptide"  
V\_region 349..672  
/product="variable region of Ig kappa light chain"

ORIGIN  
Query Match 61.1%; Score 580.6; DB 6; Length 672;  
Best Local Similarity 91.1%; Pred. No. 4.8e-144;  
Matches 644; Conservative 0; Mismatches 24; Indels 39; Gaps 1;

QY 7 GAGTCAAGCTCAGAGATGACAGGGGAGGCTTATGTCAGGCTGGAGGATCCGGAATCTC 66  
DB 1 GAGTTCACAGCTGAGTGGAGTCTGGGGAGGCTTATGTCAGGCTGGAGGATCCGGAATCTC 60  
QY 67 TCCGTGGAGCCCTGAGATTCACCTTCAGTACGTTTGAATGACATGGGTGGTCAAGGCT 126  
DB 61 TCCGTGGAGCCCTGAGATTCACCTTCAGTACGTTTGAATGACATGGGTGGTCAAGGCT 120  
QY 127 CCAGAGAGAGGGCTGGAGTGGGTGCGATATATTAGTAGTGGCAGTAGTACCATCTAT 186  
DB 121 CCAGAGAGAGGGCTGGAGTGGGTGCGATATATTAGTAGTGGCAGTAGTACCATCTAT 180  
QY 187 GGAGACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACCCGTTC 246  
DB 181 GGAGACACAGTGAAGGAGCATTCACCATCTCCAGAGCAATCCCAAGAACACCCGTTC 240  
QY 247 CTGCAATATGACAGTCTAAGTCTGAGGACACGGTCATGTATTAAGTGCAGAGATTAC 306  
DB 241 CTGCAATATGACAGTCTAAGTCTGAGGACACGGTCATGTATTAAGTGCAGAGATTAC 300  
QY 307 GGGGCTTATTGGGGCCAAAGGAGACACAGTCACTGCTCTCAGTGGAGGCGGCTCAGGC 366  
DB 301 GGGGCTTATTGGGGCCAAAGGAGACACAGTCACTGCTCTCAGTGGAGGCGGCTCAGGC 360  
QY 367 GGAGGTGGCTCTGGCGGCTGGCGGATGGGACATTGAGCTCAACCGTCTCCAGAAATCATG 426  
DB 361 GGAGGTGGCTCTGGCGGCTGGCGGATGGGACATTGAGCTCAACCGTCTCCAGAAATCATG 420  
QY 427 TCTGCATCTCCAGGGGAGAGGTCACCATGACCTGGAGTGGCAGTTCAGATGTAAGTAC 486  
DB 421 TCTGCATCTCCAGGGGAGAGGTCACCATGACCTGGAGTGGCAGTTCAGATGTAAGTAC 480  
QY 487 ATGAACGTGTTCCAAAGAGTCAAGGACACCTCCCAAAAGATGATTATGACACATCC 546  
DB 481 ATGAACGTGTTCCAAAGAGTCAAGGACACCTCCCAAAAGATGATTATGACACATCC 540  
QY 547 AAACGTCTTTTGGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGGACCTTTACTCT 606  
DB 541 AAACGTCTTTTGGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGGACCTTTACTCT 600  
QY 607 CTCACAAATGACGAGTGAAGGCTGAAGATGCTGCCACTTACTACTGCCAGAGTGAAGT 666  
DB 601 CTCACAAATGACGAGTGAAGGCTGAAGATGCTGCCACTTACTACTGCCAGAGTGAAGT 660  
QY 667 AGTAACCACTCACTGCTGGTGGTGGAGCAAGCTGAGAGTGAAGC 713  
DB 661 AGTAACCACTCACTGCTGGTGGTGGAGCAAGCTGAGAGTGAAGC 707

RESULT 15  
AR051415 672 bp DNA linear PAT 29-SEP-1999  
LOCUS AR051415  
DEFINITION Sequence 62 from patent US 5830663.  
ACCESSION AR051415  
VERSION AR051415.1 GI:5974779  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 672)  
AUTHORS Embleton,M.J., Gorochoy,G., Jones,P.T. and Winter,G.P.  
JOURNAL In situ recombinant PCR within single cells  
FEATURES  
source 1..672  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 61.1%; Score 580.6; DB 6; Length 672;  
Best Local Similarity 91.1%; Pred. No. 4.8e-144;  
Matches 644; Conservative 0; Mismatches 24; Indels 39; Gaps 1;



```
QY 7 GAGGTCAAGCTGACGAGAGTCAAGGGGAGGCTTAGTGCAGCTGAGAGGTTCCCGAAACTC 66
Db 1 GATGTGCAGCTGTGTGAGTCTGAGGAGGCTTAGTGCAGCTGAGAGGTTCCCGAAACTC 60
QY 67 TCCTGTGACGCTCTGAGATTCACTTCACTAGCTTTGAAATGCACTGGGTGTGACGGCT 126
Db 61 TCCTGTGACGCTCTGAGATTCACTTCACTAGCTTTGAAATGCACTGGGTGTGACGGCT 120
QY 127 CCAGAGAAAGGGGCTGAGAGTGGGTGCGATATATTAGTAGTGGGAGTAGTACCAATCTATAT 186
Db 121 CCAGAGAAAGGGGCTGAGAGTGGGTGCGATATATTAGTAGTGGGAGTAGTACCAATCTATAT 180
QY 187 GCAGACACAGTGAAGGAGCATTTACACATCTCCAGAGACATCCCAAGAACCTCTTTC 246
Db 181 GCAGACACAGTGAAGGAGCATTTACACATCTCCAGAGACATCCCAAGAACCTCTTTC 240
QY 247 CTGCAATGACCAAGTCTAGAGTCTGAGGACAGGTCATGTATTTACTGTGCAAGAGATTAC 306
Db 241 CTGCAATGACCAAGTCTAGAGTCTGAGGACAGGTCATGTATTTACTGTGCAAGAGATTAC 300
QY 307 GGGGCTTATTTGGGGCCAGAGGACCAAGGTCACCGTCTCCTCAGGTGAGGCGGCTCAGGC 366
Db 301 GGGGCTTATTTGGGGCCAGAGGACCTGAGTCACTGTCTCTGACAGCTAG----- 347
QY 367 GAGGTGGCTCTGGCGGCTGGCGGATCGGACATTTGAGCTCACCAGTCTTCACAGCATCATG 426
Db 348 -----CCAAATTGTTCTCACCAGTCTTCACAGCATCATG 381
QY 427 TCTGCATCTCAAGGGAGAGGGTCAACCATGACCTGAGTGCAGTTCAAGTGAAGGTAC 486
Db 382 TCTGCATCTCAAGGGAGAGGGTCAACCATGACCTGAGTGCAGTTCAAGTGAAGGTAC 441
QY 487 ATGAACTGGTTCACACAGAAGTCAAGCACCTCCCAAAAGATGAAATTTATGACACATCC 546
Db 442 ATGAACTGGTTCACACAGAAGTCAAGCACCTCCCAAAAGATGAAATTTATGACACATCC 501
QY 547 AAACGTCTTTCTGAGAGTCCCTGCTGCTTCACTGAGGAGTGGGTCTGGGACCTTTACTCT 606
Db 502 AAACGTCTTTCTGAGAGTCCCTGCTGCTTCACTGAGGAGTGGGTCTGGGACCTTTACTCT 561
QY 607 CTCACAAATCAGCAGCATGAGGCTGAAGATGCTGCACATTACTGACGAGTGGAGT 666
Db 562 CTCACAAATCAGCAGCATGAGGCTGAAGATGCTGCACATTACTGACGAGTGGAGT 621
QY 667 AGTAACCCACTGACGTTGAGTCTGGGACCAAGCTGAGCTGAAAAG 713
Db 622 AGTAATCCACTGACGTTGAGTCTGGGACCAAGCTGAGCTGAAAAG 668
```

Search completed: June 23, 2005, 09:03:47  
Job time : 6828 sec8



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 05:58:45 ; Search time 3730 Seconds  
(without alignments)  
9704.858 Million cell updates/sec

Title: US-09-403-882a-1

Perfect score: 951  
Sequence: 1 atggcgcagatcagctgcga.....ggcagaagaagcagcttag 951

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_est4:  
5: gb\_est5:  
6: gb\_est6:  
7: gb\_est7:  
8: gb\_est8:  
9: gb\_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| C 1        | 364.2 | 38.3        | 872    | 7     | CK629396 AMO-AA001 |
| C 2        | 356.8 | 37.5        | 601    | 5     | BO474958 carabid49 |
| C 3        | 354.6 | 37.3        | 1104   | 7     | CK629846 AM2-AA002 |
| C 4        | 335.4 | 35.3        | 1339   | 7     | CK629414 AM1-AA001 |
| C 5        | 334.2 | 35.1        | 1419   | 7     | CK629415 AM1-AA001 |
| C 6        | 332.6 | 35.0        | 498    | 2     | BF857861 QV1-FT020 |
| C 7        | 328.2 | 34.5        | 488    | 2     | BF857860 QV1-FT020 |
| C 8        | 324   | 34.1        | 476    | 2     | BF923394 RC0-CT004 |
| C 9        | 322.6 | 33.9        | 452    | 4     | BI055323 RC0-GN022 |
| C 10       | 311.2 | 32.7        | 585    | 4     | BI030095 IL0-MT035 |
| C 11       | 309.8 | 32.6        | 416    | 4     | BI025870 IL0-MT036 |
| C 12       | 309.8 | 32.6        | 419    | 5     | BO321312 RC0-CT047 |
| C 13       | 309.8 | 32.6        | 427    | 2     | BF752315 RC3-BN042 |
| C 14       | 309.8 | 32.6        | 437    | 5     | BO321307 RC0-CT047 |
| C 15       | 309.8 | 32.6        | 445    | 5     | BO321751 M01-CT052 |
| C 16       | 309.8 | 32.6        | 449    | 5     | BO321750 RC0-EN018 |
| C 17       | 309.8 | 32.6        | 460    | 5     | BO321747 M01-CT052 |
| C 18       | 309.8 | 32.6        | 461    | 4     | BI043087 QV1-CT017 |
| C 19       | 309.8 | 32.6        | 462    | 4     | BI030046 IL0-MT034 |
| C 20       | 309.8 | 32.6        | 464    | 4     | BI043086 QV1-CT017 |
| C 21       | 309.8 | 32.6        | 473    | 5     | BO321749 M01-CT052 |
| C 22       | 308.8 | 32.5        | 473    | 5     | BO321749 M01-CT052 |
| C 23       | 308.8 | 32.5        | 455    | 4     | BI030093 IL0-MT035 |
| C 24       | 308.8 | 32.5        | 458    | 4     | BI063593 IL3-UT011 |

|      |       |      |     |   |                    |
|------|-------|------|-----|---|--------------------|
| C 25 | 308.8 | 32.5 | 462 | 2 | BF882596 QV3-ET020 |
| C 26 | 308.8 | 32.5 | 474 | 4 | BI062949 IL3-UT011 |
| C 27 | 308.2 | 32.4 | 419 | 5 | BO380220 RC1-UT003 |
| C 28 | 308.2 | 32.4 | 424 | 5 | BO376239 M04-TN010 |
| C 29 | 308.2 | 32.4 | 459 | 4 | BI025872 IL0-MT036 |
| C 30 | 307.8 | 32.4 | 444 | 2 | BF884857 QV1-MT016 |
| C 31 | 307.2 | 32.3 | 458 | 4 | BI029994 IL0-MT034 |
| C 32 | 306.8 | 32.3 | 461 | 2 | BF811970 QV1-CT017 |
| C 33 | 306.6 | 32.2 | 455 | 4 | BI030044 IL0-MT034 |
| C 34 | 306.6 | 32.2 | 469 | 2 | BF812033 RC3-CT016 |
| C 35 | 305.8 | 32.2 | 463 | 2 | BF888263 QV1-MT022 |
| C 36 | 305   | 32.1 | 466 | 2 | BF809334 QV0-CT020 |
| C 37 | 304   | 32.0 | 456 | 5 | BO358982 M04-HT105 |
| C 38 | 303.4 | 31.9 | 653 | 5 | BO321980 RC1-NT118 |
| C 39 | 302.4 | 31.8 | 431 | 4 | BF955494 RC3-NT118 |
| C 40 | 301.8 | 31.7 | 429 | 5 | BO364984 M04-SN007 |
| C 41 | 301.4 | 31.7 | 463 | 4 | BF996225 QV1-CT019 |
| C 42 | 301.2 | 31.7 | 403 | 4 | BO375976 RC3-TN009 |
| C 43 | 301.2 | 31.7 | 439 | 5 | BO321297 QV1-CT041 |
| C 44 | 301.2 | 31.7 | 454 | 2 | BF836107 QV4-HT101 |
| C 45 | 301.2 | 31.7 | 457 | 5 | BO337457 IL0-MT035 |

## ALIGNMENTS

RESULT 1  
LOCUS CK629396/C  
DEFINITION AMO-AA0013-110902-011-H03 AA0013 Apis mellifera cDNA, mRNA  
sequence.  
ACCESSION CK629396  
VERSION CK629396.1 GI:45753871  
KEYWORDS EST.  
SOURCE Apis mellifera (honey bee)  
ORGANISM Apis mellifera  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.  
REFERENCE 1 (bases 1 to 872)  
Nunes,F.M.F., Valente,V., Sousa,J.F., Cunha,M.A.V., Pinheiro,D.G., Mala,R.M., Araujo,D.D., Costa,M.C.R., Martins,W.K., Carvalho,A.F., Monesi,N., Nascimento,A.M., Peixoto,P.M.V., Silva,M.F.R., Ramos,R.G.P., Reis,L.F.L., Dias Neto,E., Souza,S.J., Simpson,A.J.G., Zagro,M.A., Soares,A.E.E., Bitondi,M.M.G., Espresicco,E.M., Espindola,F.S., Peco-Larson,M.L., Simoes,Z.L.P., Hartfelder,K. and Silva,W.A. Jr.  
Open reading frame ESTs - an efficient strategy for analysis of the honey bee transcriptome  
Unpublished (2004)  
Contact: Silva Jr. W. A.  
Molecular Genetic and Bioinformatics Laboratory  
Department of Genetics, FMRF/USP, FUNDHERP  
Rua Tenente Catão Roxo, 2501, CEP 14051-140, Brazil  
Tel: +55 16 39639300  
Fax: +55 16 39639309  
Email: wilsonjr@usp.br  
This sequence was derived from the FAPESP Genome Program  
High quality sequence start: 66  
High quality sequence stop: 625.  
Location/Qualifiers  
1..872  
/organism="Apis mellifera"  
/mol\_type="mRNA"  
/strain="Africanized"  
/db\_xref="taxon:7460"  
/sex="female, worker"  
/dev\_stage="adult"  
/clone\_lib="AA0013"  
/note="Organ: whole body"

ORIGIN  
Query Match 38.3%; Score 364.2; DB 7; Length 872;

Best Local Similarity 74.2%; Pred. No. 3.8e-89;  
Matches 492; Conservative 0; Mismatches 155; Indels 16; Gaps 2

|            |     |   |     |
|------------|-----|---|-----|
| QY         | 72  | TGCAGCCCTGGATATTCATTTTCAGTAGCTTTGGAATGCACTGGGTTGCTGACGGCTCCAGA    | 131 |
| Db         | 763 | TGCAGACTTCTGCTACATCTTCAACAAGTTATGATATTAAGCTGGGGGAGGACGCCCTGA      | 704 |
| QY         | 132 | GAAGGGGCTGAGTAGGGGTCCGATATATTAATAGTAGGCGAGTAGTACCATCTCTATATGAGA   | 191 |
| Db         | 703 | ACAGGGACTTAGTAGTGATTTGGATGGTTTTT-TCCTGGAGAGGGGAGTACGTAAATACAAATGA | 645 |
| QY         | 192 | CACAGTGAAGGAGCATTCACCATTTCCAGAGACATATCCCAAAGAACCCCTGTTCTGCA       | 251 |
| Db         | 644 | GAAAGTTCAAGGGCAGGGCCCACTGAGTAGTATGACAAGTCTTCCACCAAGCCTATATATGA    | 585 |
| QY         | 252 | AATGACCAAGTCTAAGGTCGAGGAACAGGTCATGATTAATCTGTGCAAGAT-----          | 303 |
| Db         | 584 | GCTACTTAGGCTGACATCTGAGAGATCTTGCTGTCTTAATTTCTGTGTGAAGGGGACTACTA    | 525 |
| QY         | 304 | -----TACGGGGGCTTATTTGGGGCCAAAGGACCAAGTCAACGTCCTCTCAGGTGAGG        | 356 |
| Db         | 524 | TAGGGCTACTTTTGACTTTGGGGNCCAAAGGACCAAGTCAACGTCCTCTCATGTGAGN        | 465 |
| QY         | 357 | CGGCTCAGGGCGGAGGTGGCTCTGGCGGTGGCGGATCGGACATTAGCTTCAACCAATCTCC     | 416 |
| Db         | 464 | CGTTTCAGGGCGGAGGTGGCTCTGGGGGTGGCGGATCTGACATTTAGCTTCAACCAATCTCC    | 405 |
| QY         | 417 | AGCAATCATGTCTGTCATCTCCAGGGGAGAGGGTCAACATGACCTGACAGTGCAGTTTCAAG    | 476 |
| Db         | 404 | AGCAATCATGTCTGTCATCTCCAGGGGAGAGGGTCAACATGACCTGACAGTGCAGTTTCAAG    | 345 |
| QY         | 477 | TGTAAGGTACATGAACTGATGCCAACAGAAAGTCAGGCACTTCCCCCAAAGATGGATTTA      | 536 |
| Db         | 344 | TATACGTATACATATATTTGGTACCAACAGAAAGCTGGATCTTCCCCAGATCTCTGATTTA     | 285 |
| QY         | 537 | TGACACATCCCAAATCTGTCTTTCTGAGATCCCTGCTGCTTCAATGGCAATGGGTCTGGGAC    | 596 |
| Db         | 284 | TGACACATCCCAAAGCGGTCTCTGGAGTCCCTTTTGTCTTCAGTGCAGATGGGTCTGGGAC     | 225 |
| QY         | 597 | CTCTTACTCTCTCAACATCAGAGCATGAGGCTGAAGATGCTGACCTTACTACTATGCA        | 656 |
| Db         | 224 | CTCTTATTCTCTCAACATCAACCGAATGAGGCTGAGATGCTGCACTTATTAATGCTCA        | 165 |
| QY         | 657 | GCAAGTGAAGTAGTAACCACTCAACGTTTCGATGCTGGGACCAAGCTGAGCTGAAAACGGGC    | 716 |
| Db         | 164 | GGAGTGAAGTGGTTATTCGTACACGTTTCGAGGGGGGACCAAGCTGAGGCTGAAAACGGGC     | 105 |
| QY         | 717 | GGC 719   |     |
| Db         | 104 | GCC 102   |     |
| RESULT 2   |     |   |     |
| BQ474958/c |     |   |     |
| LOCUS      |     |   |     |
| DEFINITION |     |   |     |
| ACCESSION  |     |   |     |
| VERSION    |     |   |     |
| KEYWORDS   |     |   |     |
| SOURCE     |     |   |     |
| ORGANISM   |     |   |     |
| REFERENCE  |     |   |     |
| AUTHORS    |     |   |     |
| TITLE      |     |   |     |
| JOURNAL    |     |   |     |
| MEDLINE    |     |   |     |
| PUBMED     |     |   |     |

**COMMENT**

Contact: Foster PG  
Vogler, Entomology  
The Natural History Museum  
Cromwell Road, London, SW7 5BD  
Email: p.foster@nhm.ac.uk.

## FEATURES

Location/Qualifiers  
1. .601

ORIGIN

|                           |       |                    |            |             |
|---------------------------|-------|--------------------|------------|-------------|
| Query Match               | 37.5% | Score 356.8;       | DB 5;      | Length 601; |
| Best Local Similarity     | 76.3% | Pred. No. 3.8e-87; |            |             |
| Matches 458; Conservative | 0;    | Mismatches 127;    | Indels 15; | Gaps 1.     |

|    |     |   |     |
|----|-----|---|-----|
| OY | 134 | GGGGGCTGGAATGGGATCCGATATATATAGTGGCAGTAGTACGATCTCTATGCGAGCA    | 193 |
| Db | 600 | AGGGACTTGAATGGATTGGATGGATTTTCTCGAGAGGGGAGTACTGAATACATGACA     | 541 |
| OY | 194 | CAGTGAAGGACGATTCACTCTCCAGACAACAATCCAGAACCCCTGTTCTGCAAA        | 253 |
| Db | 540 | AGTTCAAGGGCAGGGCCACACTGATGTGACAAAGTCTCCAGCAGGCTATATGAGC       | 481 |
| OY | 254 | TGACCAGCTTAAAGTCTGAGGACACGGTCAATGATATTACTGTGCAAGAT            | 303 |
| Db | 480 | TCACTTAAGGCTACATCTGAGGACCTCTGTCTATTTCTGTGCTGAGGGGACTACTATA    | 421 |
| OY | 304 | -----TACGGGGCTTATGTGGGGCCCAAGGGACCAAGGTACACGGTCTCAGGTGAGGG    | 358 |
| Db | 420 | GGCGCTACTTGACTTGTGGGGCCAGGGACCAAGGTACCGTCTCTCTCATGTGAGGGG     | 361 |
| OY | 359 | GCTCAGGCGAGGTGGCTCTGGCGGTGGCGGATCGGACATTGAGTCAACCAGTCCAG      | 418 |
| Db | 360 | GTTCAGGGCGAGGTGGCTCTGGCGGTGGCGGATCTGACATTGAGTCAACCAGTCCAG     | 301 |
| OY | 419 | CAATCATGTCTGCATCTCCAGGGGAGAGGGTCACTGATCCTGCACTGCAAGTCAAGT     | 478 |
| Db | 300 | CAATCATGTCTGCATCTCCAGGGGAGAGGGTCACTGATCCTGCACTGCAAGTCAAGT     | 241 |
| OY | 479 | TAAAGTACATGAATGGTTCACAAGAGTCAGGCAACGCCCCCAAAAGATGGATTATG      | 538 |
| Db | 240 | TACGTTACATATATTGGTACCAAGAAAGCTGGATCTCTCCCAAGACTCTGATTTATG     | 181 |
| OY | 539 | ACACATCCAACTGTCTTCTGGAAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGACCT   | 598 |
| Db | 180 | ACACATCCAAAGTGGTCTCTGGAAGTCCCTTTTCTGCTTCAGTGGCAGTGGGTCTGGACCT | 121 |
| OY | 599 | CTTACTCTTCAATCAGACAGCATGAGGCTGAAGATGTGCACTTACTATCTGCACG       | 658 |
| Db | 120 | CTTATTCTCTCAATCAACCGAATGAGGCTGAGAGATGTGCACTTATTCTGCAAG        | 61  |
| OY | 659 | AGTGAAGAGTAAACCACTCAAGTTCGGGTCTGGGACCAAGCTGAAGCTGAAGACGGGGG   | 718 |
| Db | 60  | AGTGAAGAGTATATCCGTAACGTTTCGAGAGGGGACCAAGCTGAGCTGAAGACGGGGG    | 1   |

| RESULT 3   | LOCUS  | DEFINITION      | ACCESSION | VERSION        | KEYWORDS       | SOURCE         | ORGANISM       |
|------------|--|-----------------|-----------|----------------|----------------|----------------|----------------|
| CK629846/c | CK629846   | 1104 bp         | 1104 bp   | 1104 bp        | 1104 bp        | 1104 bp        | 1104 bp        |
|            | AM2-AA0023   | -091202-021-F02 | AA0023    | Apia mellifera | Apia mellifera | Apia mellifera | Apia mellifera |
|            |  | sequence.       |           |                |                |                |                |
|            | CK629846   |                 |           |                |                |                |                |
|            | CK629846.1   | GI:45754321     |           |                |                |                |                |
|            | EST.   |                 |           |                |                |                |                |
|            | Apia mellifera   | (honey bee)     |           |                |                |                |                |
|            | Apia mellifera   |                 |           |                |                |                |                |
|            | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prelygota;      |                 |           |                |                |                |                |
|            | Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; |                 |           |                |                |                |                |
|            | Apidae; Apia.  |                 |           |                |                |                |                |

REFERENCE  
AUTHORS  
1 (bases 1 to 1104)  
Nunes, F.M.F., Valente, V., Sousa, J.F., Cunha, M.A.V., Pinheiro, D.G.,  
Melo, R.M., Araújo, D.D., Costa, M.C.R., Martins, W.K., Carvalho, A.F.,  
Monei, N., Nascimento, A.M., Peixoto, P.M.V., Silva, M.F.R.,  
Ramos, R.G.P., Reis, L.F.L., Dias Neto, E., Souza, S.J., Simpson, A.J.G.,  
Zago, M.A., Soares, A.B.E., Bitondi, M.M.G., Espiridiao, E.M.,  
Espindola, F.S., Paco-Larson, M.L., Simoes, Z.L.P., Hartfelder, K. and  
Silva, W.A. Jr.

TITLE  
Open reading frame ESTs - an efficient strategy for analysis of the  
honey bee transcriptome

JOURNAL  
Unpublished (2004)

COMMENT  
Contact: Silva Jr, W. A.  
Molecular Genetic and Bioinformatics Laboratory  
Department of Genetics, FMRP/USP, FUNDHERP  
Rua Tenente Celso Roxo, 2501, CEP 14051-140, Brazil  
Tel: +55 16 39639300  
Fax: +55 16 39639309  
Email: wilsonjr@usp.br  
This sequence was derived from the FAPESP Genome Program  
High quality sequence start: 64  
High quality sequence stop: 594.

FEATURES  
source  
1..1104  
/organism="Apis mellifera"  
/mol\_type="mRNA"  
/strain="Africanized"  
/db\_xref="taxon:7460"  
/sex="female, worker"  
/dev\_stage="adult"  
/clone\_lib="AA0023"  
/note="Organ: whole body"

ORIGIN  
Query Match 37.3%; Score 354.6; DB 7; Length 1104;  
Best Local Similarity 74.2%; Pred. No. 1.8e-86;  
Matches 495; Conservative 0; Mismatches 154; Indels 18; Gaps 3;

76 GCCTCGATTCCTTCACTAGCTTGAATGACACTGGGTGCTGAGCTCCAG-AGAA 134  
745 GCTTCTTGCTACATCTTCACAGTTATGATATAGCTGTTGAGGACGCTTCAACAA 686  
135 GGGGCTGAGTGGGTGCGATATATAGTAGTGGCAGTAGTACATCTACTATGACAGAC 194  
685 GGGACTTGAATGATGATGATGATGATTTTCTGAGAGGAGGAGTACATGAAATGAGAA 626  
195 AGTGAAGGACGATTCACATCTCCAGAGCAATCCCAAGAACCTGTTCTGCAAT 254  
625 GTTCAAGGCGACCAACACTGATGTAGCAAGTCTCCAGACACGCTATATGAGCT 566  
255 GACCACTTGAAGTCTGAGAGCAACGCTCATATTAATGTCAGAAAGATTACG----- 308  
565 CACTAGGCTGACATCTGAGGACTCTGCTATTTTCTGCTAGAGGAGGACTATATAG 506  
309 -----GGCTTATTGGGGCCAGAGGACCAACGCTCACCGTCTCTCAGTGGAGCG 358  
505 GCGCTACTTGAATGATGAGGGGCGCAAGGACCAACGCTCACCGTCTCTCAGTGGAGCG 446  
359 GCTCAGCGGAGGTGCTTGGCGGTGGCGGATCGAATTGACTACCCAGTCTTCAG 418  
445 GTTCAGGCGGAGGTGCTTGGCGGTGGCGGATCGAATTGACTACCCAGTCTTCAG 386  
419 CATTCAATGCTGATCTCCAGAGGAGGATCAACGACTGAGTGGAGGAGTCAAGTG 478  
385 CATTCAATGCTGATCTCCAGAGGAGGATCAACGACTGAGTGGAGGAGTCAAGTG 326  
479 TAAGTACATGAATCTGTTCCACAGAGTCAAGCACTCCCAAGAAAGATGATTTATG 538  
325 TAGCTTACATATATGTTGTTACCAAGAGAGCTGATCTCCCAAGACTCTGATTTATG 266  
539 ACAATCAATCAATCTGTTCTGAGAGTCCCTGCTCGCTTCAAGTGGAGTGGTCTGGA 598  
265 ACAATCAATCAATCTGTTCTGAGAGTCCCTGCTCGCTTCAAGTGGAGTGGTCTGGA 206

Qy 599 CTTACTCTCTCAATCAACAGCAGTGAAGAGCTGAAGATGCTGCACTTACTACTCCAGC 658  
Db 205 CTTATTTCTCTCAATCAACAGCAGTGAAGAGCTGAAGATGCTGCACTTACTACTCCAGC 146  
Qy 659 AGTGAAGTGAATCAACCACTCACTTGGTGGTGGAGCCAGCTGAGCTGAAACGGAGCG 718  
Db 145 AGTGAAGTGAATCAACCACTCACTTGGTGGTGGAGCGGAGCCAGCTGAGCTGAAACGGAGCG 87

Qy 719 CCGCAGA 725  
Db 86 CCAGAAA 80

RESULT 4  
LOCUS CK629414/c 1339 bp mRNA linear EST 26-MAR-2004  
DEFINITION AM1-AA0014-041102-021-D05 AA0014 Apis mellifera cDNA, mRNA  
sequence.  
ACCESSION CK629414  
VERSION CK629414.1 GI:45753889  
KEYWORDS EST.  
SOURCE Apis mellifera (honey bee)  
ORGANISM Apis mellifera

REFERENCE  
AUTHORS  
1 (bases 1 to 1339)  
Nunes, F.M.F., Valente, V., Sousa, J.F., Cunha, M.A.V., Pinheiro, D.G.,  
Melo, R.M., Araújo, D.D., Costa, M.C.R., Martins, W.K., Carvalho, A.F.,  
Monei, N., Nascimento, A.M., Peixoto, P.M.V., Silva, M.F.R.,  
Ramos, R.G.P., Reis, L.F.L., Dias Neto, E., Souza, S.J., Simpson, A.J.G.,  
Zago, M.A., Soares, A.B.E., Bitondi, M.M.G., Espiridiao, E.M.,  
Espindola, F.S., Paco-Larson, M.L., Simoes, Z.L.P., Hartfelder, K. and  
Silva, W.A. Jr.

TITLE  
Open reading frame ESTs - an efficient strategy for analysis of the  
honey bee transcriptome

JOURNAL  
Unpublished (2004)

COMMENT  
Contact: Silva Jr, W. A.  
Molecular Genetic and Bioinformatics Laboratory  
Department of Genetics, FMRP/USP, FUNDHERP  
Rua Tenente Celso Roxo, 2501, CEP 14051-140, Brazil  
Tel: +55 16 39639300  
Fax: +55 16 39639309  
Email: wilsonjr@usp.br  
This sequence was derived from the FAPESP Genome Program  
High quality sequence start: 62  
High quality sequence stop: 662.

FEATURES  
source  
1..1339  
/organism="Apis mellifera"  
/mol\_type="mRNA"  
/strain="Africanized"  
/db\_xref="taxon:7460"  
/sex="female, worker"  
/dev\_stage="adult"  
/clone\_lib="AA0014"  
/note="Organ: whole body"

ORIGIN  
Query Match 35.3%; Score 335.4; DB 7; Length 1339;  
Best Local Similarity 80.3%; Pred. No. 3.9e-81;  
Matches 412; Conservative 0; Mismatches 86; Indels 15; Gaps 1;

222 AGACATCCCAAGAACACCTGTTCTGCAAAATGACCAATGATGATGATGATGATGATGAT 281  
Db 597 AGACATCCCTCCAGACACAGCTATATGAGCTCACTAGGCTCAATGATGATGATGATGAT 538  
Qy 282 CATGATTTACTGTCAGAG-----ATTAGGGGCTTATTTAGGGGCAAG 326  
Db 537 TGTCTATTTCTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 478  
Qy 327 GACACGCTGACCGTCTCTCAGGTGAGCGGCTCAGGCGGAGGAGGAGGAGGAGGAGGAGG 386

|    |     |   |     |
|----|-----|---|-----|
| Db | 477 | GACCAAGGTACCGTCTCTCATGTCGAGGCGGTTCAAGCGGAGGTTGGCTCTGGGGTGG    | 418 |
| Qy | 387 | CGAATGGACATTTAGATCTACCCCACTCTCCAGCAATCATGTCTGCATCTCCAGGGAGAG  | 446 |
| Db | 417 | CGGATCTGCACATTGAGCTCACCACAGTCTCCAGCAATCATGTCTGCATCTCCAGGGAGAG | 358 |
| Qy | 447 | GGTACCATGACCTGCAGTGGCAGTTTCAAGGTAAAGTACATGAATCGTGTCCACAGAA    | 506 |
| Db | 357 | GGTACCATGACCTGCAGTGGCAGTTCAAAGTAAAGTACATGAATCGTGTCCACAGAA     | 298 |
| Qy | 507 | GTCAAGCACCTCCCCCAAAAGATGATTTATGACATCCAAACGTCTTTCGAGTCCC       | 566 |
| Db | 297 | GCCTGGATCTCTCCCCAGACTCTGATTTATGACATCCAAACGTCTTTCGAGTCCC       | 238 |
| Qy | 567 | TGCTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTTCACAAATCAGCAGATGA    | 626 |
| Db | 237 | TTTTTCGCTTCAGTGGCAGTGGGTCTGGGACCTCTTACTCTTCACAAATCAGCAGATGA   | 178 |
| Qy | 627 | GGCTGAAGATCTCTGCCACTTACTCTGCCAGCAGTGGAGTAACTCCACTACGTTGGG     | 686 |
| Db | 177 | GGCTGAAGATCTCTGCCACTTACTCTGCCAGCAGTGGAGTAACTCCACTACGTTGGG     | 118 |
| Qy | 687 | TGCTGGGACCAAGCTGGAGCTGAAACGGGCGGC                             | 719 |
| Db | 117 | AGGGGGACCAAGCTGGAGCTGAAACGGGCTGCC                             | 85  |

| RESULT 5   | CK629415/c   | 1419 bp | mRNA | linear | EST 26-MAR-2004 |
|------------|--|---------|------|--------|-----------------|
| LOCUS      | CK629415   |         |      |        |                 |
| DEFINITION | AA1-AA0014-041102-021-H09 AA0014 Apis mellifera cDNA, mRNA sequence. |         |      |        |                 |

|          |                            |
|----------|----------------------------|
| SOURCE   | Apis mellifera (honey bee) |
| ORGANISM | Apis mellifera             |

REFERENCE  
AUTHORS  
1 (bases 1 to 1419)  
Nunes, F.M.F., Valente, V., Sousa, J.F., Cunha, M.A.V., Pinheiro, D.G.,

| TITLE   |
|---|
| Open reading frame ESTs - an efficient strategy for analysis of the honey bee transcriptome |

**JOURNAL COMMENT**  
Unpublished (2004)  
Contact: Silva Jr, W. A.  
Molecular Genetic and Bioinformatics Laboratory  
Department of Genetics, FMPP/USP, FUNDHERP  
Rua Tenente Catao Roxo, 2501, CEP 14051-140, Brazil

```

This sequence was derived from the FAPESP Genome Program
High quality sequence start: 63
High quality sequence stop: 605.
Location/Qualifiers
1..1419
FEATURES
source
```

```

/organism="Apis mellifera"
/mol_type="RNA"
/strain="Africanized"
/db_xref="taxon:7460"
/sex="female, worker"
/dev_stage="adult"
/clone_lib="Aa0014"
/note="Organ: whole body"

```

|                       |       |              |         |            |     |        |      |
|-----------------------|-------|--------------|---------|------------|-----|--------|------|
| Query Match           | 35.1% | Score        | 334.2   | DB         | 7   | Length | 1419 |
| Best Local Similarity | 75.7% | Pred. No.    | 8.5e-81 |            |     |        |      |
| Matches               | 433   | Conservative | 0       | Mismatches | 124 | Indels | 15   |
|                       |       |              |         |            |     | Gaps   | 1    |

|    |     |  |     |
|----|-----|--|-----|
| Qy | 157 | ATTAGTAGTGCGAGTAGTACATCTACTATATGACAGACAGAGMAAGGGACGATTCACATC     | 216 |
|    |     |  |     |
| Db | 662 | ATTTTTCCTGAGAGGGGANGTACTGAAATCAATGAGAAGTTCAAGGGGAGGCACACTG       | 603 |
|    |     |  |     |
| Qy | 217 | TCCAGAGACAAATCCCAAGAACACCCTTTCTCTGCAATGACCACTTAAGTCTGAGAC        | 276 |
|    |     |  |     |
| Db | 602 | AGTGTAGACAAAGTCTCCACAGACAGCCTATATGAGAGTCACTAGCTGACACTGAGAC       | 543 |
|    |     |  |     |
| Qy | 277 | ACGCTATGTATTACTGTGCAAGAGAT-----TACGGGGCTTATTGGGGC                | 321 |
|    |     |  |     |
| Db | 542 | TCGTCTGTCTATTTTCTGTCTGTAAGGGGAGTACTATAGCGGCTACTTTGACTTTGTGGGC    | 483 |
|    |     |  |     |
| Qy | 322 | CAAGGAGCCACGGTCAACCGTCTCTCAGTGTGAGAGCGGCTCAGGCGGAGGTGCTTGGC      | 381 |
|    |     |  |     |
| Db | 482 | CAAGGGGACACGGGTCAACCGTCTCTCCTCATGTGAGAGACGGTTACGGCGGAGGTGGCTGTGC | 423 |
|    |     |  |     |
| Qy | 382 | GGTGGGGGAGACCGGACATTGAGCTCACCCGACTCCACCAATCAATGTCGATCTGCCAAGG    | 441 |
|    |     |  |     |
| Db | 422 | GGTGGCGGATCTCAATGAGCTCACCCGACTCCACCAATCAATGTCGATCTGCCAAGG        | 363 |
|    |     |  |     |
| Qy | 442 | GAGAGGGTCAACATGACCTGACAGTGCAGTTCAAAGTGAAGGTACATGAATCGTTCCAA      | 501 |
|    |     |  |     |
| Db | 362 | GAGAGGGTCAACATGACCTGACAGTGCAGTTCAAAGTGAAGGTACATGAATCGTTCCAA      | 303 |
|    |     |  |     |
| Qy | 502 | CAGAAATCAGGCACTCTCCGCCAAAAGATGATTTATGACACATCCAAACTGTCTTCTGGA     | 561 |
|    |     |  |     |
| Db | 302 | CAGAAAGCTGGAACTCTCCGCCAACTCTTATTTATGACACATCCAAAGTGTCTTCTGGA      | 243 |
|    |     |  |     |
| Qy | 562 | GTCCTCTGCTGCTTCAATGAGCAGTGGGGTCTGGGACCTCTTACTCTTCACAAATGAGAGC    | 621 |
|    |     |  |     |
| Db | 242 | GTCCTCTTTCGCTTCAATGAGCAGTGGGTCTGGGACCTCTTATTTCTTCACAAATGAGAGC    | 183 |
|    |     |  |     |
| Qy | 622 | ATGAGAGCTGAAGATGTGCACTTACTACTGCGACGATGAGTAGTAACTCACTCAG          | 681 |
|    |     |  |     |
| Db | 182 | ATGAGAGCTGAAGATGTGCACTTATTTACTGCGACGAGATGAGATGTCTTATCCGTACAG     | 123 |
|    |     |  |     |
| Qy | 682 | TTTGGTGTCTGGGACCAAGCTGAGAGCTGAACG                                | 713 |
|    |     |  |     |
| Db | 122 | TTTGGAGGGGGGAGCCAAAGCTGAGAGCTGAACG                               | 91  |

|            |  |
|------------|--|
| RESULT     | 6  |
| BFB87861/c |  |
| LOCUS      | BFB87861          498 bp        mRNA     linear    EST 16-JAN-2001 |
| DEFINITION | OVI-FR0202-071100-465-g03 FR0202 Homo sapiens cDNA, mRNA sequence. |
| ACCESSION  | BFB87861   |
| VERSION    | BFB87861.1   GI:12245605   |
| KEYWORDS   | EST.   |
| SOURCE     | Homo sapiens (human)   |
| ORGANISM   | Homo sapiens   |

| REFERENCE          | AUTHORS   | TITLE  | JOURNAL                       | MEDLINE | PUBMED           | COMMENT |
|--------------------|---|--|-------------------------------|---------|------------------|---------|
| 1 (bases 1 to 498) | Dias Nero, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H., Brentani, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J. | Shotgun sequencing of the human transcriptome with ORF expressed sequence tags | Proc. Natl. Acad. Sci. U.S.A. | 97 (7)  | 3491-3496 (2000) |         |
|                    | Contact: Simpson A.J.G.   |  |                               |         | 10737800         |         |

Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP  
Brazil

Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV1&t2=QV1-FT0202-071100-465-f03&t3=2000-11-07&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 8  
 High quality sequence stop: 497.  
 Location/Qualifiers

## FEATURES

source

1..498  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_11b="FT0202"  
 /note="Organ: prostate tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 35.0%; Score 332.6; DB 2; Length 498;  
 Best Local Similarity 88.1%; Pred. No. 1.8e-80;  
 Matches 362; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 316 TGGGGCCAGAGGACCAAGGTACCGTCTCTCAGGTGAGAGCGGCTCAGCGAGGTGCG 375  
 DB 495 TGGGGCCAGAGGACCAAGGTACCGTCTCTCAGGTGAGAGCGGCTCAGCGAGGTGCG 436  
 QY 376 TTTGGCGGTGGGGGATCGAGCATTTGAGTCAACCCATCTCCAGCATCATGTCTGATCT 435  
 DB 435 TCTGGCGGTGGGGGATCGAGCATTTGAGTCAACCCATCTCCAGCATCATGTCTGATCT 376  
 QY 436 CCAGGGGAGAGGGGTCAAGTCACTGAGTGCAGTTCAGTGAAGTGAATGATGATGAGTGA 495  
 DB 375 CCAGGGGAGAGGGGTCAAGTCACTGAGTGCAGTTCAGTGAAGTGAATGATGATGAGTGA 316  
 QY 496 TTCCAACAGAGTCAAGGACCTTCCCAAGATGATTATGACATCCAAATGTCT 555  
 DB 315 TACCAACAGAGTCAAGGACCTTCCCAAGATGATTATGACATCCAAATGTCT 256  
 QY 556 TTGTAGTCTCTGCTGCTTCAAGTGCAGTGGGTCTGGGACCTCTTACTCTCTCAATTC 615  
 DB 255 CTTGGAGTCTCTTCTTCTTCAAGTGCAGTGGGTCTGGGACCTCTTACTCTCTCAATTC 196  
 QY 616 AGCAGCATGAGGCTGAAGATGTGCTTACTTACTGACAGCATGAGTGAATGATGAGTGA 675  
 DB 195 AACCGATGAGGCTGAAGATGTGCTTACTTACTGACAGCATGAGTGAATGATGAGTGA 136  
 QY 676 CTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726  
 DB 135 TACAGCTTGGAGGGGGGACCAAGCTGAGTGAATGAGTGAATGAGTGAATGAGTGA 85

RESULT 7  
 BF857860/c 488 bp mRNA linear EST 16-JUN-2001  
 LOCUS QV1-FT0202-071100-465-f03 FT0202 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF857860  
 ACCESSION BF857860  
 VERSION BF857860.1 GI:12245604  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 488)  
 Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT  
 Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV1&t2=QV1-FT0202-071100-465-f03&t3=2000-11-07&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 8  
 High quality sequence stop: 487.  
 Location/Qualifiers

## FEATURES

source

1..488  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_11b="FT0202"  
 /note="Organ: prostate tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 34.5%; Score 328.2; DB 2; Length 488;  
 Best Local Similarity 88.1%; Pred. No. 2.9e-79;  
 Matches 357; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 322 CAAGGACCAAGGATCAAGTCTCTCAGGTGAGAGCGGCTCAGCGAGGTGCTCTGCG 381  
 DB 488 CAAGGACCAAGGATCAAGTCTCTCAGGTGAGAGCGGCTCAGCGAGGTGCTCTGCG 429  
 QY 382 GGTGGGATCGGACATTGAGTCAACCCAGTCTCCAGCATCATGTCTGATCTCCAGGG 441  
 DB 428 GATGGGATCGGACATTGAGTCAACCCAGTCTCCAGCATCATGTCTGATCTCCAGGG 369  
 QY 442 GAGAGGTTACCATGACCTGCGAGTGCAGTTCAGTGAAGTGAATGAGTGAATGAGTGA 501  
 DB 368 GAGAGGTTACCATGACCTGCGAGTGCAGTTCAGTGAAGTGAATGAGTGAATGAGTGA 309  
 QY 502 CAGAAGTCAAGGACCTTCCCAAGATGATTATGACATCCAAATGTCTTGGG 561  
 DB 308 CAGAAGTCAAGGACCTTCCCAAGATGATTATGACATCCAAATGTCTTGGG 249  
 QY 562 GTCCCTGCTGCTTCAAGTGCAGTGGGTCTGGGACCTCTTACTCTTCAATCAGCAGC 621  
 DB 248 GTCCCTGCTGCTTCAAGTGCAGTGGGTCTGGGACCTCTTACTCTTCAATCAGCAGC 189  
 QY 622 ATGAGAGCTGAAGATGCTGCTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661  
 DB 188 ATGAGAGCTGAAGATGCTGCTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129  
 QY 682 TTGCGTGTGGGACCAAGCTGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 726  
 DB 128 TTGCGTGTGGGACCAAGCTGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 84

RESULT 8  
BF923394/c 476 bp mRNA linear EST 19-JUN-2001  
LOCUS QV4-NT0248-271100-585-c10 NT0248 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF923394  
ACCESSION BF923394.1 GI:12319282  
VERSION EST.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 476)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
COMMENT Contact: Simpson A.J.J.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV4&f2=QV4-NT0248-  
271100-585-c10&t3=2000-11-27&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 42  
High quality sequence stop: 395.  
Location/Qualifiers  
1..476  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NT0248"  
/note="Organ: nervous tumor; Vector: puc18; Site 1: Sma1;  
Site 2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN  
Query Match 34.1%; Score 324; DB 2; Length 476;  
Best Local Similarity 87.6%; Pred. No. 4,1e-78;  
Matches 354; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 316 TGGGCGCAGGAGACACCGCTACCGTCTCCAGGTGGAGGGCGTCCAGGAGAGTGGC 375  
DB 466 TGGGCGCAGGAGACACCGCTACCGTCTCCAGGTGGAGGGCGTCCAGGAGAGTGGC 407  
QY 376 TCTGGCGGTGGCGATCGGACATTGAGCTCACCAGTCTCCAGCATCATGTCTGCATCT 435  
DB 406 TCTGTGGGTGGCGATCGGACATTGAGCTCACCAGTCTCCAGCATCATGTCTGCATCT 347  
QY 436 CCAAGGGAGAGGGGTACCATGACCTGCACTGGCACTTCAAGTGTAAAGTACATGAACCTG 495  
DB 346 CCAAGGGAGAGGGGTACCATGACCTGCACTGGCACTTCAAGTGTAAAGTACATGAACCTG 287  
QY 496 TTCCAAAGAAAGTACGAGCACTCCCAAAAGATGATTATGACACATCCAAAGTGTCT 555  
DB 286 TACCAACGAAAGCTGATCTCTCCCAAGATCTCGATTATGACACATCCAAAGTGTCT 227

QY 556 TCTGAGTCCCTGCTCGCTTACAGTGCAGTGGGTCTGGAGACCTTACTCTCTCAATC 615  
DB 226 CTTGAGTCCCTGCTCGCTTACAGTGCAGTGGGTCTGGAGACCTTACTCTCTCAATC 167  
QY 616 AGCAGCATGAGAGCTGAAGATGCTGCCCTTACTCTGCTCCAGCATGAGTGTAACTCA 675  
DB 166 TACCGAATGAGAGCTGAAGATGCTGCCCTTACTCTGCTCCAGCATGAGTGTAACTCA 107  
QY 676 CTCACGTTCCGAGTGGAGACCAAGCTGAGCTGAACAGCGAGCGC 719  
DB 106 TACACGTTCCGAGTGGAGACCAAGCTGAGCTGAACAGCGAGCGC 63

RESULT 9  
BI055323/c 452 bp mRNA linear EST 15-JUN-2001  
LOCUS RCO-GN0233-060201-031-b08 GN0233 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BI055323  
ACCESSION BI055323.1 GI:14462853  
VERSION EST.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 452)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
COMMENT Contact: Simpson A.J.J.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RCO&f2=RCO-GN0233-  
060201-031-b08&t3=2001-02-06&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 42  
High quality sequence stop: 452.  
Location/Qualifiers  
1..452  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="GN0233"  
/note="Organ: placenta normal; Vector: puc18; Site 1:  
Sma1; Site 2: Sma1; A mini-library was made by cloning  
products derived from ORESTES PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the puc 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

ORIGIN  
Query Match 33.9%; Score 322.6; DB 4; Length 452;  
Best Local Similarity 88.8%; Pred. No. 9.9e-78;  
Matches 349; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 334 GTCAACGCTCTCTCAAGTGGAGGGCGCTCAGGCGAGGTGCTTGGCGGTGGCGAGTGC 393

Db 452 GTCACCGCTCTCCATGAGGCGGTTCCAGGGAGGTGGCTCTGGCGGATCT 393  
Oy 394 GACATTGAGCTACCCAGTCTCCAGCAATCATGTCTGCAGGGAGAGGCTCAC 453  
Db 392 GACATTGAGCTACCCAGTCTCCAGCAATCATGTCTGCAGGGAGAGGCTCAC 333  
Oy 454 ATGACCTGAGTGCAGTTCAGTCTAAGTATCATGAATGCTGTTCCAGAGTCA 513  
Db 332 ATGACCTGAGTGCAGTTCAGTCTAAGTATCATGAATGCTGTTCCAGAGTCA 273  
Oy 514 ACCTCCCAAGATGATTTATGACATCCAACTGTCTTGTGAGTCTCTGCTCG 573  
Db 272 TCTCTCCCAAGATGATTTATGACATCCAACTGTCTTGTGAGTCTCTGCTCG 213  
Oy 574 TTCAGTGGAGTGGTCTGGGACCTCTTCTCTGACATGAGCATGAGGCTGAA 633  
Db 212 TTCAGTGGAGTGGTCTGGGACCTCTTCTCTGACATGAGCATGAGGCTGAA 153  
Oy 634 GATGCTGCACCTTACTACTGCGCAGAGTGAAGTAACTCACTGCTGCTGAG 693  
Db 152 GATGCTGCACCTTACTACTGCGCAGAGTGAAGTAACTCACTGCTGCTGAG 93  
Oy 694 ACCAAGCTGAGTGAACGGGCGCGCGAGAA 726  
Db 92 ACCAAGCTGAGTGAACGGGCGCGCGAGAA 60

RESULT 10  
B1030095 585 bp mRNA linear EST 14-JUN-2001  
LOCUS IL0-MT0355-220301-500-d11 MT0355 Homo sapiens cDNA, mRNA sequence.  
DEFINITION B1030095  
VERSION B1030095.1 GI:14436725  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 585)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&cl=IL0-MT0355-  
220301-500-d11&cl3=2001-03-22&cl4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 28  
High quality sequence stop: 584.  
Location/Qualifiers  
1..585  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_11b="MT0355"

/note="Organ: marrow; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

ORIGIN  
Query Match 32.7%; Score 311.2; DB 4; Length 585;  
Best Local Similarity 85.8%; Pred. No. 1.5e-74;  
Matches 357; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

Oy 304 TACGGGGCTTATTTGGGGCCAGGAGACAGGTCACCGTCTCTCAGTGAAGCGGCTCA 363  
Db 503 TACTTTGACTTTGGGCCCCAGGAGACAGGTCACCGTCTCTCTTAATGTGAAGCGGTTCA 444  
Oy 364 GGGGAGGTGGCTCTGGCGGTGGCGGATGACATTGAAGTCAACCAAGTCTCAGCAATC 423  
Db 443 GCGGAGTGGCTCTGACCGGTGGCGGATGACATTGAAGTCAACCAAGTCTCAGCAATC 384  
Oy 424 ATGTGCAATCTCCAGGGGAGAGGTCACCAATGACCTGAGTGCAGTCAAGTGAAG 483  
Db 383 ATGTGCAATCTCCAGGGGAGAGGTCACCAATGACCTGAGTGCAGTCAAGTGAAG 324  
Oy 484 TACATGAATCTGTTCCAGAGAGTCAAGGACCTCCCAAGATGATTTATGACACA 543  
Db 323 TACATGATTTGGATGACCAAGAGAGTCAAGGACCTCCCAAGATGATTTATGACACA 264  
Oy 544 TCCAAATCTGTTCTTGGAGTCCCTGCTCGCTTCAAGTGAAGTGGTCTTGGAGCTTTAC 603  
Db 263 TCCAAATCTGTTCTTGGAGTCCCTGCTCGCTTCAAGTGAAGTGGTCTTGGAGCTTTAC 204  
Oy 604 TCTCTCAATCAAGCAGATGAGGCTGAAGATGCTGCACTTACTGACGACAGTGG 663  
Db 203 TCTCTCAATCAAGCAGATGAGGCTGAAGATGCTGCACTTACTGACGACAGTGG 144  
Oy 664 AGTAGTAACCACTCACTGCTGGTGGTGGAGCAAGTGAAGGCGGCGC 719  
Db 143 AGTAGTAACCACTCACTGCTGGTGGTGGAGCAAGTGAAGGCGGCGC 89

RESULT 11  
B1025870 416 bp mRNA linear EST 14-JUN-2001  
LOCUS IL0-MT0364-050201-500-g01 MT0364 Homo sapiens cDNA, mRNA sequence.  
DEFINITION B1025870  
VERSION B1025870.1 GI:14432500  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 416)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br



This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL0&t2=IL0-MT0364-050201-500-501&t3=2001-02-05&t4=1>)

Seq primer: puc 18 forward  
High quality sequence start: 18  
High quality sequence stop: 416.  
Location/Qualifiers

FEATURES  
Source

```
1..416
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /dev_stage="Adult"
  /clone_idb="MT0364"
  /note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
```

## ORIGIN

Query Match 32.6%; Score 309.8; DB 4; Length 416;  
Best Local Similarity 88.9%; Pred. No. 3.3e-74;  
Matches 335; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
350 GTGAGGCGGCTCAGCGGAGGTGCTCTGCGCGTGGCGATCGACATTTGAGCTCACC 409
|||||
415 GTGAGGCGGCTCAGCGGAGGTGCTCTGCGCGTGGCGATCGACATTTGAGCTCACC 356
|||||
410 AGTCTCCAGCAATCATGTCTGATCTCCAGGGAGAGGGTCAACATGACCTGACAGTCCA 469
|||||
355 AGTCTCCAGCAATCATGTCTGATCTCCAGGGAGAGGGTCAACATGACCTGACAGTCCA 296
|||||
470 GTTCAAGTGAAGTACATGAAGTCTGCTTCCACAGAACTGACGACCTCCCAAAAGAT 529
|||||
295 GCTCAAGTGAAGTACATGAAGTCTGCTTCCACAGAACTGACGACCTCCCAAAAGAT 236
|||||
530 GGATTTATGACATCCAACTGTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 589
|||||
225 TGATTTATGACATCCAACTGTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 176
|||||
590 CTGGGACCTTCTTCTGCTTCTGACATGACAGCATGAGGCTGAGATGCTGCACCTTACT 649
|||||
175 CTGGGACCTTCTTCTGCTTCTGACATGACAGCATGAGGCTGAGATGCTGCACCTTACT 116
|||||
650 ACTGCACAGCATGAGTGAATCAACCTGAGTGGTCTGAGACCAAGCTGAGCTGA 709
|||||
115 ACTGCACAGCATGAGTGAATCAACCTGAGTGGTCTGAGACCAAGCTGAGCTGA 56
|||||
710 AACGGGCGGCGCGCAGAA 726
|||||
55 AACGGGCGGCGCGCAGAA 39
|||||
```

RESULT 12  
BQ321312/c 419 bp mRNA linear EST 17-MAY-2002  
LOCUS RC0-CT0472-080800-031-f01 CT0472 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BQ321312  
ACCESSION BQ321312  
VERSION BQ321312.1 GI:20929288  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 419)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brijones, M. R.,  
Nagai, M. A., da Silva, M. Jr., Zago, M. A., Bordin, S., Costa, F. F.,  
Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,  
Brunstein, A. J., de Oliveira, P. S., Bucher, P., Jongeneel, C. V.,  
O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and

TITLE Simpson, A. J.  
JOURNAL Shotgun sequencing of the human transcriptome with ORF expressed  
PROC. NATL. ACAD. SCI. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC0&t2=RC0-CT0472-080800-031-f01&t3=2000-08-08&t4=1>)

FEATURES  
Source

```
1..419
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /dev_stage="Adult"
  /clone_idb="CT0472"
  /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
```

## ORIGIN

Query Match 32.6%; Score 309.8; DB 5; Length 419;  
Best Local Similarity 88.9%; Pred. No. 3.3e-74;  
Matches 335; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
350 GTGAGGCGGCTCAGCGGAGGTGCTCTGCGCGTGGCGATCGACATTTGAGCTCACC 409
|||||
418 GTGAGGCGGCTCAGCGGAGGTGCTCTGCGCGTGGCGATCGACATTTGAGCTCACC 359
|||||
410 AGTCTCCAGCAATCATGTCTGATCTCCAGGGAGAGGGTCAACATGACCTGACAGTCCA 469
|||||
358 AGTCTCCAGCAATCATGTCTGATCTCCAGGGAGAGGGTCAACATGACCTGACAGTCCA 299
|||||
470 GTTCAAGTGAAGTACATGAAGTCTGCTTCCACAGAACTGACGACCTCCCAAAAGAT 529
|||||
298 GCTCAAGTGAAGTACATGAAGTCTGCTTCCACAGAACTGACGACCTCCCAAAAGAT 239
|||||
530 GGATTTATGACATCCAACTGTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 589
|||||
228 TGATTTATGACATCCAACTGTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 179
|||||
590 CTGGGACCTTCTTCTGCTTCTGACATGACAGCATGAGGCTGAGATGCTGCACCTTACT 649
|||||
178 CTGGGACCTTCTTCTGCTTCTGACATGACAGCATGAGGCTGAGATGCTGCACCTTACT 119
|||||
650 ACTGCACAGCATGAGTGAATCAACCTGAGTGGTCTGAGACCAAGCTGAGCTGA 709
|||||
118 ACTGCACAGCATGAGTGAATCAACCTGAGTGGTCTGAGACCAAGCTGAGCTGA 59
|||||
710 AACGGGCGGCGCGCAGAA 726
|||||
58 AACGGGCGGCGCGCAGAA 42
|||||
```

RESULT 13  
BQ752315/c 427 bp mRNA linear EST 10-JAN-2001  
LOCUS BQ752315



DEFINITION RC3-BN0428-201100-011-f06 BN0428 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF752315  
 VERSION BF752315.1 GI:12078991  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 427)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 JOURNAL MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC3&t2=RC3-BN0428-201100-011-f06&t3=2000-11-20&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 34  
 High quality sequence stop: 427.  
 Location/Qualifiers  
 1..427  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="BN0428"  
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

## FEATURES

source

Query Match 32.6%; Score 309.8; DB 2; Length 427;  
 Best Local Similarity 88.9%; Pred. No. 3.4e-74;  
 Matches 335; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

350 GTGAGAGCGGCTCAGCGGAGGAGTGGCTGGCGGATGGAGCATTTAGCTACCC 409  
 |||||  
 426 GTGAGAGCGGCTCAGCGGAGGAGTGGCTGGCGGATGGAGCATTTAGCTACCC 357  
 |||||  
 410 AGTCTCAGCAATCATGTCTGCATCTCCAGGGAGAGGGTCAACATGACTCGTAGTCCA 469  
 |||||  
 366 AGTCTCAGCAATCATGTCTGCATCTCCAGGGAGAGGGTCAACATGACTCGTAGTCCA 307  
 |||||  
 470 GTTCAAGGTGAAGTCAATGAACTGGTTCCACAGAAATGACGACCTCCCCCAAAAGAT 529  
 |||||  
 306 GGTCAAGGTGAAGTCAATGAACTGGTTCCACAGAAATGACGACCTCCCCCAAAAGAT 247  
 |||||  
 530 GGATTTATGACATCAATCAAACTGTCTTGGAGAGTCCCTGCTCAGTGGAGAGTGGGT 589  
 |||||  
 246 TGAATTTATGACATCAATCAAACTGTCTTGGAGAGTCCCTGCTCAGTGGAGAGTGGGT 187  
 |||||  
 590 GTGGAGCTTAACTCTCTCAATCAGACGATGAGAGGCTGAAGTGTCCACTTACT 649  
 |||||

Db 186 GTGGAGCTTAACTCTCTCAATCAGACCCGAATGAGAGTGTCCACTTAACT 127  
 QY 650 ACTGACGAGAGTGAAGTAACTTAACCCATGACGTTGGTCTGGAGACCAAGTGAAGCTGA 709  
 |||||  
 Db 126 ACTGACGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 67  
 |||||  
 QY 710 AACGGCGGCGCGCAGAA 726  
 |||||  
 Db 66 AACGGCGGCGCGCAGAA 50  
 |||||

RESULT 14  
 BQ321307/c 437 bp mRNA linear EST 17-MAY-2002  
 LOCUS RC0-CT0472-080800-031-d01 CT0472 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BQ321307  
 ACCESSION BQ321307  
 VERSION BQ321307.1 GI:20929277  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 437)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 JOURNAL MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC0&t2=RC0-CT0472-080800-031-d01&t3=2000-08-08&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 38  
 High quality sequence stop: 437.  
 Location/Qualifiers  
 1..437  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0472"  
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

TITLE  
 JOURNAL MEDLINE 20202663  
 PUBMED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC0&t2=RC0-CT0472-080800-031-d01&t3=2000-08-08&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 38  
 High quality sequence stop: 437.  
 Location/Qualifiers  
 1..437  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0472"  
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## FEATURES

source

Query Match 32.6%; Score 309.8; DB 5; Length 437;  
 Best Local Similarity 88.9%; Pred. No. 3.4e-74;  
 Matches 335; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 350 GTGAGAGCGGCTCAGCGGAGGAGTGGCTGGCGGATGGAGCATTTAGCTACCC 409  
 |||||  
 Db 436 GTGAGAGCGGCTCAGCGGAGGAGTGGCTGGCGGATGGAGCATTTAGCTACCC 377  
 |||||

## ORIGIN

QY 410 AGTCTCCAGCAATCATGTTCTGATCTCCAGAGGAGAGGGTCCATCATGACCTGAGTCCCA 469  
DB 376 AGTCTCCAGCAATCATGTTCTGATCTCCAGAGGAGAGGGTCCATCATGACCTGAGTCCCA 317  
QY 470 GTTCAAGTAAAGTATGATGAACTGGTCTCAACAGAAAGTCAAGGACCTCCCCAAAGAT 529  
DB 316 GCTCAAGTAAAGTATGATGAACTGGTCTCAACAGAAAGTCAAGGACCTCCCCAAAGAT 257  
QY 530 GGATTATGACATCATCAAACTGTTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 589  
DB 256 TGATTATATACATCATCAAGTGGCTCTGAGTCCCTTCTGCTTCACTGAGTGGGT 197  
QY 590 CTGGAGCCCTTACTCTCTCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649  
DB 196 CTGGAGCCCTTACTCTCTCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 137  
QY 650 ACTGCCAGAGTGAAGTAACTCACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 709  
DB 136 ACTGCCAGAGTGAAGTAACTCACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 77  
QY 710 AACGGGCGGCGCAGAA 726  
DB 76 AACGGGCGGCGCAGAA 60

RESULT 15  
BQ321751/c 445 bp mRNA linear EST 17-MAY-2002  
LOCUS MR1-CT0529-250900-003-d01 CT0529 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BQ321751  
ACCESSION BQ321751  
VERSION BQ321751.1 GI:20930355  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 445) Eukaryota; Primates; Catarrhini; Homiidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

FEATURES  
source Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR1&2=MR1-CT0529-  
250900-003-d01&3=2000-09-25&4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 22  
High quality sequence stop: 445.  
Location/Qualifiers  
1..445  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="CT0529"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.

196,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

ORIGIN  
Query Match 32.6%; Score 309.8; DB 5; Length 445;  
Best Local Similarity 88.9%; Pred. No. 3.4e-74;  
Matches 335; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
QY 350 GTGAGAGGAGCTCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409  
DB 444 GTGAGAGGAGCTTCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 385  
QY 410 AGTCTCCAGCAATCATGTTCTGATCTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 469  
DB 384 AGTCTCCAGCAATCATGTTCTGATCTCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 325  
QY 470 GTTCAAGTAAAGTATGATGAACTGGTCTCAACAGAAAGTCAAGGACCTCCCCAAAGAT 529  
DB 324 GCTCAAGTAAAGTATGATGAACTGGTCTCAACAGAAAGTCAAGGACCTCCCCAAAGAT 265  
QY 530 GGATTATGACATCATCAAACTGTTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 589  
DB 264 TGATTATGACATCATCAAACTGTTCTTCTGAGTCCCTGCTGCTTCACTGAGTGGGT 205  
QY 590 CTGGAGCCCTTACTCTCTCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649  
DB 204 CTGGAGCCCTTACTCTCTCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 145  
QY 650 ACTGCCAGAGTGAAGTAACTCACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 709  
DB 144 ACTGCCAGAGTGAAGTAACTCACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 85  
QY 710 AACGGGCGGCGCAGAA 726  
DB 84 AACGGGCGGCGCAGAA 68

Search completed: June 23, 2005, 10:06:04  
Job time : 3735 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 05:54:00 ; Search time 613 seconds  
(without alignments)  
9183.808 Million cell updates/sec

Title: US-09-403-882A-1  
Perfect score: 951  
Sequence: 1 atggcgaaggtcaagctgca.....ggcagaagaagccacgttag 951

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn19808:\*  
2: geneseqn19908:\*  
3: geneseqn20008:\*  
4: geneseqn20018:\*  
5: geneseqn20028:\*  
6: geneseqn20038:\*  
7: geneseqn20048:\*  
8: geneseqn20058:\*  
9: geneseqn20068:\*  
10: geneseqn20078:\*  
11: geneseqn20088:\*  
12: geneseqn20098:\*  
13: geneseqn20108:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 951   | 100.0       | 951    | 2     | AAZ20266    |
| 2          | 949   | 99.8        | 8902   | 4     | AAH78812    |
| 3          | 934.4 | 98.3        | 6115   | 2     | AAH78812    |
| 4          | 749.6 | 78.8        | 1256   | 3     | AAH78812    |
| 5          | 749.6 | 78.8        | 1256   | 4     | AAH78812    |
| 6          | 749.6 | 78.8        | 1256   | 4     | AAH78812    |
| 7          | 749.6 | 78.8        | 1256   | 4     | AAH78812    |
| 8          | 749.6 | 78.8        | 1256   | 4     | AAH78812    |
| 9          | 699.4 | 73.5        | 726    | 6     | AAH78812    |
| 10         | 655.4 | 68.9        | 738    | 2     | AAH78812    |
| 11         | 651   | 68.5        | 711    | 2     | AAH78812    |
| 12         | 580.6 | 61.1        | 672    | 2     | AAH78812    |
| 13         | 548   | 57.6        | 1725   | 2     | AAH78812    |
| 14         | 539   | 56.7        | 1668   | 2     | AAH78812    |
| 15         | 532.4 | 56.0        | 711    | 9     | AAH78812    |
| 16         | 510   | 53.6        | 738    | 10    | AAH78812    |
| 17         | 510   | 53.6        | 1157   | 10    | AAH78812    |
| 18         | 474.4 | 49.4        | 774    | 2     | AAH78812    |
| 19         | 470   | 49.4        | 841    | 10    | AAH78812    |
| 20         | 468.6 | 49.3        | 807    | 3     | AAH78812    |

|    |       |      |      |    |          |                    |
|----|-------|------|------|----|----------|--------------------|
| 21 | 467   | 49.1 | 927  | 5  | AAH78812 | AAH78812 DNA encod |
| 22 | 465   | 48.9 | 888  | 4  | AAH78812 | AAH78812 DNA encod |
| 23 | 462.8 | 48.7 | 708  | 2  | AAH78812 | AAH78812 DNA encod |
| 24 | 462.2 | 48.6 | 870  | 2  | AAH78812 | AAH78812 DNA encod |
| 25 | 461.2 | 48.5 | 840  | 3  | AAH78812 | AAH78812 DNA encod |
| 26 | 460.6 | 48.4 | 807  | 3  | AAH78812 | AAH78812 DNA encod |
| 27 | 460.2 | 48.4 | 807  | 3  | AAH78812 | AAH78812 DNA encod |
| 28 | 460.2 | 48.4 | 807  | 3  | AAH78812 | AAH78812 DNA encod |
| 29 | 459.6 | 48.3 | 873  | 12 | AAH78812 | AAH78812 DNA encod |
| 30 | 453.2 | 47.7 | 702  | 2  | AAH78812 | AAH78812 DNA encod |
| 31 | 452   | 47.5 | 756  | 12 | AAH78812 | AAH78812 DNA encod |
| 32 | 449.6 | 47.3 | 1679 | 2  | AAH78812 | AAH78812 DNA encod |
| 33 | 449.6 | 47.3 | 1679 | 2  | AAH78812 | AAH78812 DNA encod |
| 34 | 449.6 | 47.3 | 1679 | 5  | AAH78812 | AAH78812 DNA encod |
| 35 | 449.6 | 47.3 | 1679 | 5  | AAH78812 | AAH78812 DNA encod |
| 36 | 448.8 | 47.2 | 726  | 8  | AAH78812 | AAH78812 DNA encod |
| 37 | 448.8 | 47.2 | 726  | 13 | AAH78812 | AAH78812 DNA encod |
| 38 | 446   | 46.9 | 759  | 2  | AAH78812 | AAH78812 DNA encod |
| 39 | 446   | 46.9 | 861  | 12 | AAH78812 | AAH78812 DNA encod |
| 40 | 445.4 | 46.8 | 810  | 13 | AAH78812 | AAH78812 DNA encod |
| 41 | 444.2 | 46.7 | 717  | 2  | AAH78812 | AAH78812 DNA encod |
| 42 | 442   | 46.5 | 717  | 4  | AAH78812 | AAH78812 DNA encod |
| 43 | 442   | 46.5 | 717  | 10 | AAH78812 | AAH78812 DNA encod |
| 44 | 442   | 46.5 | 717  | 12 | AAH78812 | AAH78812 DNA encod |
| 45 | 438.2 | 46.1 | 720  | 6  | AAH78812 | AAH78812 DNA encod |

## ALIGNMENTS

|          |  |                                  |
|----------|--|----------------------------------|
| RESULT 1 | AAZ20266   | AAZ20266 standard; cDNA; 951 BP. |
| ID       | AAZ20266   |                                  |
| XX       | AAZ20266   |                                  |
| AC       | AAZ20266   |                                  |
| DT       | 17-JAN-2000 (first entry)  |                                  |
| XX       |  |                                  |
| DE       | cDNA coding single chain antibody used in probe detection.   |                                  |
| XX       |  |                                  |
| KW       | Single chain antibody; scAb; scFv; spectroscopic probe; ss.  |                                  |
| XX       |  |                                  |
| OS       | Unidentified.  |                                  |
| XX       |  |                                  |
| PN       | MO9951986-A1.  |                                  |
| XX       |  |                                  |
| PD       | 14-OCT-1999.   |                                  |
| XX       |  |                                  |
| PF       | 08-APR-1999; 99WO-US007847.  |                                  |
| XX       |  |                                  |
| PR       | 08-APR-1998; 98US-0081118P.  |                                  |
| PR       | 09-APR-1998; 98US-0081340P.  |                                  |
| XX       |  |                                  |
| PA       | (REGC) UNIV CALIFORNIA.  |                                  |
| XX       |  |                                  |
| PI       | Farinas J;   |                                  |
| XX       |  |                                  |
| DR       | WPI; 1999-611066/52.   |                                  |
| XX       |  |                                  |
| PT       | P-PSDB; AAJ32086.  |                                  |
| PT       | Localizing probes to specific sites in cells that express single-chain antibody reactive with probe-ligand conjugate, particularly for detecting post-translational modification and its modulators.   |                                  |
| XX       |  |                                  |
| PS       | Disclosure; Page 47-48; 69pp; English.   |                                  |
| XX       |  |                                  |
| CC       | This cDNA codes for a single chain antibody (scAb, see AAJ32086) that has 2 c-myc epitopes. The cDNA can be amplified from plasmid pHook-1 by PCR (see also AAZ20267-68). The scAb is used in claimed methods for localizing a probe and for detecting a post-translational activity in a cell, and is expressed by a claimed transgenic non-human animal. The invention provides methods and reagents for targeting probes to selected cellular locations through the expression of specific binding partners |                                  |

CC (such as scAbs) within a cell, and for creating assays for post-translation activities. The invention allows the monitoring of the location of such intracellular, specific binding partners over time and in response to stimuli, such as test chemicals. Spectroscopic probes can be used to screen a test chemical for activity. The invention also includes cells and transgenic organisms comprising the intracellular specific binding partner, where the specific binding partner can bind with the spectroscopic probe/ligand conjugate. scAbs can be expressed within the cell and can be designed to bind a wide variety of spectroscopic probes, including small molecules that have better (and more diverse) spectroscopic properties than green fluorescent protein

Sequence 951 BP; 226 A; 247 C; 272 G; 206 T; 0 U; 0 Other;

Query Match 100.0%; Score 951; DB 2; Length 951;  
Best Local Similarity 100.0%; Pred. No. 1.1e-242;  
Matches 951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCGAGGTCAAGTGCAGAGAGGAGGCTTATGTCAGCTTGGAGGCTCCGG 60  
DB 1 ATGGCCGAGGTCAAGTGCAGAGAGGAGGCTTATGTCAGCTTGGAGGCTCCGG 60  
QY 61 AAATCTCTGTGCAAGCTCTGGAATTCATTTCAGTAGCTTGAATGCACTGGGTGGT 120  
DB 61 AAATCTCTGTGCAAGCTCTGGAATTCATTTCAGTAGCTTGAATGCACTGGGTGGT 120  
QY 121 CAGGCTCCAGAGAGAGGAGGCTGAGTGGGTGCATATATATATAGTAGGCAAGTACATC 180  
DB 121 CAGGCTCCAGAGAGAGGAGGCTGAGTGGGTGCATATATATAGTAGGCAAGTACATC 180  
QY 181 TACTATGACAGACAGTGAAGGAGAGATTCACATCTCCAGAGACAAATCCCAAGAACCC 240  
DB 181 TACTATGACAGACAGTGAAGGAGAGATTCACATCTCCAGAGACAAATCCCAAGAACCC 240  
QY 241 CTGTTCTCTGCAATGACAGCTCTAAGGTCTGAGAGACAGCGTATGTAATCTGTCAGAGA 300  
DB 241 CTGTTCTCTGCAATGACAGCTCTAAGGTCTGAGAGACAGCGTATGTAATCTGTCAGAGA 300  
QY 301 GATTACGGGGCTTTATTTGGGGCCAAAGGACACAGCTGCACCTGCTCTCAGTGGAGGGCGC 360  
DB 301 GATTACGGGGCTTTATTTGGGGCCAAAGGACACAGCTGCACCTGCTCTCAGTGGAGGGCGC 360  
QY 361 TCAGGGGAGAGGAGGCTGCGGCTGGCGGATGGACATTAAGCTACCCAGTCTCAGAGA 420  
DB 361 TCAGGGGAGAGGAGGCTGCGGCTGGCGGATGGACATTAAGCTACCCAGTCTCAGAGA 420  
QY 421 ATCATGTCTGCATCTCCAGGGAGAGGGTCACTACATGACCTGAGTGCAGTTCAAGTGA 480  
DB 421 ATCATGTCTGCATCTCCAGGGAGAGGGTCACTACATGACCTGAGTGCAGTTCAAGTGA 480  
QY 481 AGGTACATGAACTGTTCCCAACAGAAAGTCAAGGACCTCCCAAAAGATGATTTATGAC 540  
DB 481 AGGTACATGAACTGTTCCCAACAGAAAGTCAAGGACCTCCCAAAAGATGATTTATGAC 540  
QY 541 ACATCCAACTCTCTTTGGAGAGTCCCTGCTCCCTTCAAGAGAGTGGGTCTGGGACCTCT 600  
DB 541 ACATCCAACTCTCTTTGGAGAGTCCCTGCTCCCTTCAAGAGAGTGGGTCTGGGACCTCT 600  
QY 601 TACTCTCTCAATAGCAGCAGATGAGAGGTGAAGATGTCGACATTAATCTAGTCCAGAG 660  
DB 601 TACTCTCTCAATAGCAGCAGATGAGAGGTGAAGATGTCGACATTAATCTAGTCCAGAG 660  
QY 661 TGGAGTAGTAACTCACTACGTTCCGGTGTGGGACCAAGCTGGAATCGGAGCGGCGCC 720  
DB 661 TGGAGTAGTAACTCACTACGTTCCGGTGTGGGACCAAGCTGGAATCGGAGCGGCGCC 720  
QY 721 GGAGAAACAAAACCTATCTCAGAAAGAGATCTGAATGGGGCCGTGACAAACAAAACCTC 780  
DB 721 GGAGAAACAAAACCTATCTCAGAAAGAGATCTGAATGGGGCCGTGACAAACAAAACCTC 780  
QY 781 ATCTCAGAAAGAGATCTGATCTGTGGGCGCAGAGACGACGAGAGAGTCTATCTGTGTCGA 840  
DB 781 ATCTCAGAAAGAGATCTGATCTGTGGGCGCAGAGACGACGAGAGAGTCTATCTGTGTCGA 840

QY 841 CACTCTTGCCCTTTAAGGTGGTGTGATCTCAGCCATCTTGCCCTTGGTGTGCTACCC 900  
DB 841 CACTCTTGCCCTTTAAGGTGGTGTGATCTCAGCCATCTTGCCCTTGGTGTGCTACCC 900  
QY 901 ATCATCTCCCTTATCATCTCATCATCTGTTTGGCAGAAAGCCACGTTAG 951  
DB 901 ATCATCTCCCTTATCATCTCATCATCTGTTTGGCAGAAAGCCACGTTAG 951

## RESULT 2

AAH78812  
ID AAH78812 standard; DNA; 8902 BP.

XX AAH78812;

DT 03-MAY-2002 (first entry)

XX Candidate peptide inhibitor expressing plasmid vector (pBLOCK-3-8).

XX Reverse n-hybrid screening method; pBLOCK-3.8 plasmid vector; de;

XX biological antagonist; biological inhibitor; peptide inhibitor;

XX common binding partner; leukaemia; solid tumour.

XX Synthetic.

XX WO20016787-A1.

XX 13-SEP-2001.

XX 08-MAR-2001; 2001WO-US007669.

XX 08-MAR-2000; 2000AU-00006131.

XX 23-MAR-2000; 2000AU-00006437.

XX 11-APR-2000; 2000AU-00006830.

XX 06-NOV-2000; 2000AU-00001256.

XX (TWNT-) TWN TELETHON INST CHILD HEALTH RES.

XX Hopkins R, Serebriiskii I, Watt PM, Golemis E;

XX WPI; 2001-582279/65.

XX Claim 47; Page 93-98; 125pp; English.

XX The invention comprises a reverse n-hybrid screening method for  
CC identifying antagonists or inhibitors of biological interactions. The  
CC method involves the use of multiple genes to distinguish antagonists or  
CC inhibitors of one interaction from those of other interactions. The  
CC method identifies peptides which inhibit a target interaction (between at  
CC least two binding partners), but does not inhibit a non-target  
CC interaction within a host cell. The method further comprises a series of  
CC plasmid vectors designed to express candidate peptide inhibitors. The  
CC method of the invention has enhanced specificity in the detection of  
CC interacting binding partners compared to conventional reverse n-hybrid  
CC screens. The method of the invention also has the capacity to distinguish  
CC between interactions that involve one or more common binding partners.  
CC The method of the invention is useful for identifying a peptide which  
CC inhibits a target interaction. The peptides identified by the method of  
CC the invention may be used in the treatment of leukaemia and solid  
CC tumours. The present sequence represents a plasmid vector (pBLOCK-3.8) of  
CC the invention that is designed to express a candidate peptide inhibitor  
XX

Sequence 8902 BP; 2200 A; 2201 C; 2134 G; 2365 T; 0 U; 2 Other;

Query Match 99.8%; Score 949; DB 4; Length 8902;  
Best Local Similarity 99.8%; Pred. No. 8.2e-242;  
Matches 949; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

|    |   |               |
|----|---|---------------|
| QY | ATGACCGAGGCTCAAGCTGCAGAGATCAGGGGAGAGCTTAAGTCAGACCTCGAGAGGGTCTCCGG | 60            |
| Db | ATGGCCGAGGCTCAAGCTGCAGAGATCAGGGGAGAGCTTAAGTCAGACCTCGAGAGGGTCTCCGG | 2820 2879     |
| QY | AAACTCTCTCTGAGAGCTCTGGAATTCATCTTCACTAGTAGCTTGGAAATGCACTGGGATTCGT  | 120           |
| Db | AAACTCTCTCTGAGAGCTCTGGAATTCATCTTCACTAGTAGCTTGGAAATGCACTGGGATTCGT  | 2880 2939     |
| QY | CAGGCTCCAGAGAAAGGGGCTGAGTGGGTCCCATATATTAGTAGTGCATGATGCATC         | 180           |
| Db | CAGGCTCCAGAGAAAGGGGCTGAGTGGGTCCCATATATTAGTAGTGCATGATGCATC         | 121 2940 2999 |
| QY | TACTATGTCAGACACAGTGAAGGACGATTCACATCTCCAGAGCAATCCCAAGAACCC         | 240           |
| Db | TACTATGTCAGACACAGTGAAGGACGATTCACATCTCCAGAGCAATCCCAAGAACCC         | 181 3000 3059 |
| QY | CTGTTCCTGCAAAATGACACGATCTAAGTCTGAGGACACGGTCACTATTTACTGTGCAGA      | 300           |
| Db | CTGTTCCTGCAAAATGACACGATCTAAGTCTGAGGACACGGTCACTATTTACTGTGCAGA      | 241 3060 3119 |
| QY | GATTACGGGGCTTATTGGGGCCAAAGGACACACGGTCAACGGTCTCTCAGGTGAGCGGC       | 360           |
| Db | GATTACGGGGCTTATTGGGGCCAAAGGACACACGGTCAACGGTCTCTCAGGTGAGCGGC       | 301 3120 3179 |
| QY | TCAGGCGGAGGTGGCTCTGGGGGGTGGCGGATTCGACATTTGAGCTCACCCACTCTCCAGCA    | 420           |
| Db | TCAGGCGGAGGTGGCTCTGGGGGGTGGCGGATTCGACATTTGAGCTCACCCACTCTCCAGCA    | 361 3180 3239 |
| QY | ATCATATGTCGATCTCCAGGGGAGAGGGTCAACATGACCTGACAGTGCAGTTCAAGTGT       | 480           |
| Db | ATCATATGTCGATCTCCAGGGGAGAGGGTCAACATGACCTGACAGTGCAGTTCAAGTGT       | 421 3240 3299 |
| QY | AGGTATCATGAACCTGGTTCACAAGAGTCAGGCACCTCCCAAAAGATGATTTATGAC         | 540           |
| Db | AGGTATCATGAACCTGGTTCACAAGAGTCAGGCACCTCCCAAAAGATGATTTATGAC         | 481 3300 3359 |
| QY | ACATCCCAAACGTCTTCTGAGAGTCCCTGCTGCTTCAATGAGAGGGGTCTGGGACCTCT       | 600           |
| Db | ACATCCCAAACGTCTTCTGAGAGTCCCTGCTGCTTCAATGAGAGGGGTCTGGGACCTCT       | 541 3360 3419 |
| QY | TACTCTCTCAAAATCAGCAGATGGAAGGTGCTGACACTTACTGTGCAGAG                | 660           |
| Db | TACTCTCTCAAAATCAGCAGATGGAAGGTGCTGACACTTACTGTGCAGAG                | 601 3420 3479 |
| QY | TGAGTAGTAACCACTCAGCTTCGGTCTGAGACCAAGCTGAGACTGAAACGGGCGGC          | 720           |
| Db | TGAGTAGTAACCACTCAGCTTCGGTCTGAGACCAAGCTGAGACTGAAACGGGCGGC          | 661 3480 3539 |
| QY | GCAGAACAAAACCTCATCTCAGAAAGAGATCTGAATGGGGCCGTGAGAACAAAATCTC        | 780           |
| Db | GCAGAACAAAACCTCATCTCAGAAAGAGATCTGAATGGGGCCGTGAGAACAAAATCTC        | 721 3540 3599 |
| QY | ATCTCAGAAAGAGATCTGAATGCTGTGGGCCAGGACACGACAGAGGTCACTGCTGGTGC       | 840           |
| Db | ATCTCAGAAAGAGATCTGAATGCTGTGGGCCAGGACACGACAGAGGTCACTGCTGGTGC       | 781 3600 3659 |
| QY | CACCTCTTGCCCTTTAAGTGGTGTGATCTCAGCCATCTCGCCCTGCTGCTGCTCAC          | 900           |
| Db | CACCTCTTGCCCTTTAAGTGGTGTGATCTCAGCCATCTCGCCCTGCTGCTGCTCAC          | 841 3660 3719 |
| QY | ATCATATCTCCCTTATCATCTCATCATCTGTTGGCAGAAAGAACCCAGTTAG              | 951           |
| Db | ATCATATCTCCCTTATCATCTCATCATCTGTTGGCAGAAAGAACCCAGTTAG              | 901 3720 3770 |

|          |                                 |
|----------|---------------------------------|
| RESULT 3 |                                 |
| AAT63235 |                                 |
| ID       | AAT63235 standard; DNA; 6115 BP |
| XX       |                                 |
| AC       | AAT63235;                       |
| XX       |                                 |
| DT       | 27-AUG-2003 (revised)           |

| DT | 18-JUN-1997   | (first entry)   |
|----|---|---|
| XX | Eukaryotic expression vector plasmid pPhox.TM.                    |   |
| XX | Plasmid pPhox.TM; eukaryotic expression vector; transfected cell; |   |
| KW | single chain antibody; sFv; molecular hook; ds.                   |   |
| XX | Cytomegalovirus.  |   |
| OS | unidentified bacteriophage; T7.                                   |   |
| OS | Mus; sp.  |   |
| OS | influenza virus;  |   |
| OS | Bos; taurus.  |   |
| OS | Homo; sapiens.  |   |
| OS | Escherichia; coli.  |   |
| OS | unidentified bacteriophage; SP6.                                  |   |
| OS | Synthetic.  |   |
| OS | Chimeric.   |   |
| XX | Key   |   |
| FT | promoter  | Location/Qualifiers                                       |
| FT |   | 1. .596   |
| FT | /tag= a   | /note= "cytomegalovirus immediate early promoter permits  |
| FT | /note= "high level expression of the sFv in eukaryotic cells"     |   |
| FT | 638. .657   |   |
| FT | promoter  |   |
| FT |   | /tag= b   |
| FT | /note= "T7 promoter"  |   |
| FT | 737. .1795  |   |
| FT | /tag= c   |   |
| FT | 737. .799   |   |
| FT | /tag= d   | /note= "murine Ig kappa-chain V-J2-C signal peptide       |
| FT | /note= "directs the sFv to the plasma membrane for extracellular  |   |
| FT | display"  |   |
| FT | 800. .826   |   |
| FT | /tag= e   | /product= "haemagglutinin A epitope tag"                  |
| FT | /note= "allows detection of the sFv by Mab 12CA5"                 |   |
| FT | 842. .1555  |   |
| FT | /tag= f   | /product= "pPhox sFv"                                     |
| FT | /note= "pPhox sFv single chain antibody recognises the            |   |
| FT | happen phox and allows isolation or detection of cells            |   |
| FT | displaying this sFv"  |   |
| FT | 1568. .1600   |   |
| FT | /tag= g   | /product= "Myc.1 epitope 1"                               |
| FT | /note= "allows detection of sFv by Mab 9E10.2"                    |   |
| FT | 1613. .1645   |   |
| FT | /tag= h   | /product= "myc.1 epitope 2"                               |
| FT | /note= "allows detection of sFv by Mab 9E10.2"                    |   |
| FT | 1646. .1795   |   |
| FT | /tag= i   | /product= "PDGFR transmembrane domain"                    |
| FT | /note= "fusion of platelet derived growth factor receptor         |   |
| FT | transmembrane domain to sFv anchors the antibody to the           |   |
| FT | plasma membrane for display"                                      |   |
| FT | 1831. .1848   |   |
| FT | /tag= j   | /note= "Sf6 promoter"                                     |
| FT | 1853. .2081   |   |
| FT | /tag= k   | /note= "bovine growth hormone polyA signal permits proper |
| FT | /note= "processing and polyadenylation of the mRNA for            |   |
| FT | stabilisation of the message"                                     |   |
| FT | 2212. .2795   |   |
| FT | /tag= l   | /function= "Col E1 origin"                                |
| FT | /note= "Col E1 origin allows high copy replication and            |   |
| FT | growth in E. coli"  |   |
| FT | complement(2980. .3251)   |   |
| FT | /tag= o   | /note= "thymidine kinase polyA site"                      |
| FT | /note= "thymidine kinase polyA site"                              |   |

```

FT CDS complement(3426..4214)
FT /tag= n
FT /note= "neomycin/kanamycin resistance gene allows
FT selection of the plasmid in E. coli and confers G418
FT resistance to mammalian cells"
FT complement(4249..4587)
FT /tag= m
FT /function= "SV40 origin and promoter"
FT CDS complement(4666..5526)
FT /tag= p
FT /note= "ampicillin resistance gene allows selection of
FT the plasmid in E. coli"
FT /tag= .6113
FT /function= "f1 origin"
FT misc_feature
FT /tag= q
FT /function= "f1 origin"
XX MO9708186-A1.
XX
XX 06-MAR-1997.
XX
XX 23-AUG-1996; 96WO-US015819.
XX
XX 24-AUG-1995; 95US-00518835.
XX
XX (INV1-) INVTROGEN CORP.
XX
XX Cheanut RD, Baytan A, Hoefler JP, Bernhard A, Chang M;
XX
XX WPI; 1997-179169/16.
XX
XX Eukaryotic expression vector - for identification and separation of
XX transfected cells from total cell population.
XX
XX Example 1; Fig 6; 83pp; English.
XX
XX Bkaryotic expression vector plasmid pPhox-TM (PAT63235) can be used in a
XX novel system for isolating and identifying transfected eukaryotic cells.
XX It employs a single chain antibody (scFv) directed against a hapten, phox,
XX to isolate transiently transfected cells from total populations in
XX culture. The expressed fusion protein, phox scFv, also contains 2 epitope
XX tag peptides and the transmembrane domain (TM) of human platelet derived
XX growth factor receptor (PDGFR). When expressed in transfected cells, the
XX fusion protein is anchored to the membrane via the PDGFR-TM, and the
XX antibody binding unit, phox scFv, is exposed to the extracellular
XX environment. The transfected cells are selected using antigen (phox)-
XX coated magnetic beads. The method allows for identification and selection
XX of transfected cells as early as 2 hr after transfection, thus allowing
XX study of the acute effects of the expression of the gene of interest.
XX (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 6115 BP; 1440 A; 1607 C; 1599 G; 1467 T; 0 U; 2 Other;
XX
XX Query Match 98.3%; Score 934.4; DB 2; Length 6115;
XX Best Local Similarity 99.4%; Pred. No. 5.5e-238;
XX Matches 948; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
XX
XX 1 ATGCCGAGGTCAAGCTGAGAGTCAAGGAGGAGCTTATGTCAGCTTGAAGGTCCTCCG 60
XX 842 ATGGCGGAGGTCAAGCTGAGAGTCAAGGAGGAGCTTATGTCAGCTTGAAGGTCCTCCG 901
XX
XX 61 AAATCTCTCTGTCAGAGCTCTGATTCACCTTCAGTAGCTTTGGAATGACAGTGGTTCT 120
XX
XX 902 AAATCTCTCTGTCAGAGCTCTGATTCACCTTCAGTAGCTTTGGAATGACAGTGGTTCT 961
XX
XX 121 CAGGCTCCAGAGAGGGCTGAGTGGTGGCATATATTAGTAGTGGCAGTAGTACCATC 180
XX
XX 962 CAGGCTCCAGAGAGGGCTGAGTGGTGGCATATATTAGTAGTGGCAGTAGTACCATC 1021
XX
XX 181 TACTATGACAGACAGTGAAGGAGCAGATTCACCATCTCCAGAGACATCCCAAGAACACC 240
XX
XX 1022 TACTATGACAGACAGTGAAGGAGCAGATTCACCATCTCCAGAGACATCCCAAGAACACC 1081
XX
XX 241 CTGTTCTGCAAAATGACCACTAGCTGAGAGACAGGCTATGATTATCTGTGCAAGA 300

```

```

DB 1082 CTGTTCTGCAAAATGACCACTAGCTGAGAGCAGGAGCATGATTATCTGTGCAAGA 1141
XX
XX 301 GATTACGGGGCTTATTTGGGGCCCAAGGAGCAGGTCACCGTCTCTCAGTGGAGGCGGC 360
XX
XX 1142 GATTACGGGGCTTATTTGGGGCCCAAGGAGCAGGTCACCGTCTCTCAGTGGAGGCGGC 1201
XX
XX 361 TCAGGCGAGAGTGTCTGTCGGTGGTGGCGATTCGACATTTAGCTCACCAGTCTTCAGCA 420
XX
XX 1202 TCAGGCGAGAGTGTCTGTCGGTGGTGGCGATTCGACATTTAGCTCACCAGTCTTCAGCA 1261
XX
XX 421 ATCATGTCTGCAATCTCCAGGGGAGAGGGTCAACATGACCTTGAGTCCAGTTCAAGTGA 480
XX
XX 1262 ATCATGTCTGCAATCTCCAGGGGAGAGGGTCAACATGACCTTGAGTCCAGTTCAAGTGA 1321
XX
XX 481 AGGTACATGAACTGTTCCAAAGAGTCAAGGACCTCCCCCAAAAGATGATTATGAC 540
XX
XX 1322 AGGTACATGAACTGTTCCAAAGAGTCAAGGACCTCCCCCAAAAGATGATTATGAC 1381
XX
XX 541 ACATCCAAACTGTCTTCTGAGAGTCCCTGCTGCTTCACTGAGCAGTGGTCTGGGACTTCT 600
XX
XX 1382 ACATCCAAACTGTCTTCTGAGAGTCCCTGCTGCTTCACTGAGCAGTGGTCTGGGACTTCT 1441
XX
XX 601 TACTCTTCACAATGAGAGAGATGAGAGCTGAAGATGCTGCCACTTACTACTGCCAGCAG 660
XX
XX 1442 TACTCTTCACAATGAGAGAGATGAGAGCTGAAGATGCTGCCACTTACTACTGCCAGCAG 1501
XX
XX 661 TGAAGTAGTAACCCACTGACGTTCCGTTCTGGGAGCAGGCTGAGAGCTGAAACGG--GCG 717
XX
XX 1502 TGAAGTAGTAACCCACTGACGTTCCGTTCTGGGAGCAGGCTGAGAGCTGAAACGG--GCG 1561
XX
XX 718 GCCGAGAACCAAAATCATCTTCAGAGAGATCTGAATGGGCGCTGAGCAGCAACAAA 777
XX
XX 1562 GCCGAGAACCAAAATCATCTTCAGAGAGATCTGAATGGGCGCTGAGCAGCAACAAA 1621
XX
XX 778 CTCATCTCAGAGAGAGATCTGAATGGGCGCTGAGCAGCAGGATCATCTGGTGG 837
XX
XX 1622 CTCATCTCAGAGAGAGATCTGAATGGGCGCTGAGCAGCAGGATCATCTGGTGG 1681
XX
XX 838 CCACACTCTTGGCCCTTTAAGGTGGTGTGATCTTCAGCCATCTGGCCCTGGTGGTGTCTC 897
XX
XX 1682 CCACACTCTTGGCCCTTTAAGGTGGTGTGATCTTCAGCCATCTGGCCCTGGTGGTGTCTC 1741
XX
XX 898 ACCATCATCTCCCTTATCATCTCTCATCTGCTTTGGCAGAAAGACCAAGCTTAA 951
XX
XX 1742 ACCATCATCTCCCTTATCATCTCTCATCTGCTTTGGCAGAAAGACCAAGCTTAA 1795
XX
XX
XX RESULT 4
XX AAC66074
XX ID AAC66074 standard; DNA; 1256 BP.
XX
XX AC AAC66074;
XX
XX DT 22-FEB-2001 (first entry)
XX
XX XX
XX DE E. coli expression plasmid pUBS520-ScFvOx DNA.
XX
XX XX Eukaryotic protein; protease; interferon; antibody; hormone;
XX
XX KM disulfide bridge; de.
XX
XX OS Escherichia coli.
XX
XX XX
XX PN EP1048732-A1.
XX
XX PD 02-NOV-2000.
XX
XX XX
XX PF 26-APR-1999; 99EP-00107412.
XX
XX PR 26-APR-1999; 99EP-00107412.
XX
XX XX
XX PA (HOPF ) HOPFMANN LA ROCHE & CO AG F.

```

XX MPI; 2000-674185/66.  
 DR P-PSDB; AAB11398.  
 XX  
 XX  
 PT Preparation of water-soluble eukaryotic polypeptides with disulfide  
 bridges e.g. rPA, comprises cultivation of prokaryotic cells in the  
 presence of arginine or amide compound.  
 PT  
 XX  
 PS Example 6; Page 20-22; 40pp; German.

CC This invention describes a novel preparation of a water-soluble,  
 CC naturally occurring eukaryotic polypeptide containing two or more  
 CC cysteine units bound via a disulfide bridge which comprises cultivation  
 CC of prokaryotic cells in the presence of arginine or an amide compound.  
 CC The method is useful for the preparation of eukaryotic proteins e.g.  
 CC proteases, interferons, protein hormones, antibodies or antibody  
 CC fragments (e.g. a single chain FV fragment that binds to thyroid  
 CC stimulating hormone). It is especially useful for preparing proteins with  
 CC more than five disulfide bridges, e.g. recombinant plasminogen activator  
 CC (rPA). The technique is simple and does not require in vitro after-  
 CC treatment, such as the removal of inclusion bodies, reduction or  
 CC naturalization.

SQ Sequence 1256 BP; 304 A; 320 C; 339 G; 293 T; 0 U; 0 Other;

Query Match 78.8%; Score 749.6; DB 3; Length 1256;

Best Local Similarity 98.2%; Pred. No. 4.8e-189;

Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGGCCGAGGTCAAGTGCAGAGAGTCAAGGAGGCTTAGTGAGGCTTGAGGCTCCCG 60  
 DB 199 ATGGCCGAGGTCAAGTGCAGAGAGTCTGGGGAGGCTTAGTGAGGCTTGAGGCTCCCG 258  
 QY 61 AAACCTCTCTGGGACCTCTGGATTCACCTTCAGTACGCTTGGAATGACCTGGGTGCT 120  
 DB 259 AAACCTCTCTGGGACCTCTGGATTCACCTTCAGTACGCTTGGAATGACCTGGGTGCT 318  
 QY 121 CAGGCTCCAGAGAGGGGCTGGAGTGGTGCATATATTAGTAGGAGTACATC 180  
 DB 319 CAGGCTCCAGAGAGGGGCTGGAGTGGTGCATATATTAGTAGGAGTACATC 378  
 QY 181 TACTATGACAGACAGATGAAGGAGCATTCACCATCTCCAGAGCAATCCAGAACCC 240  
 DB 379 TACTATGACAGACAGATGAAGGAGCATTCACCATCTCCAGAGCAATCCAGAACCC 438  
 QY 241 CTGTTCTCTCAATATCCAGTCTTAAGTCTGAGGACACGGTCTATATTACTGTGAAGA 300  
 DB 439 CTGTTCTCTCAATATCCAGTCTTAAGTCTGAGGACACGGTCTATATTACTGTGAAGA 498  
 QY 301 GATTACGGGGCTTATGGGGCCCAAGGAGCAACGGTCAACGGTCTCCAGGTGAGGCGGC 360  
 DB 499 GATTACGGGGCTTATGGGGCCCAAGGAGCAACGGTCAACGGTCTCCAGGTGAGGCGGC 558  
 QY 361 TCAGGGGAGGTGGCTCTGGCGGTGGCGGATCGGACATTTGAGCTCAACCATCTCCAGA 420  
 DB 559 TCAGGGGAGGTGGCTCTGGCGGTGGCGGATCGGACATTTGAGCTCAACCATCTCCAGA 618  
 QY 421 ATCATGTCTGATCTCTCAAGGGAGAGGGTCAACATGACTGACAGTCCAGTTCAAGTGA 480  
 DB 619 ATCATGTCTGATCTCTCAAGGGAGAGGGTCAACATGACTGACAGTCCAGTTCAAGTGA 678  
 QY 481 AGGTACATGAATCGTTCCAAAGAGTACAGGACCTCCCGCCAAAGATGATTTATGAC 540  
 DB 679 AGGTACATGAATCGTTCCAAAGAGTACAGGACCTCCCGCCAAAGATGATTTATGAC 738  
 QY 541 ACATCCAAATCTGTTCTGGAGTCCCTGTCTGCTTCACTGAGTGGAGTCTGGGACCTCT 600  
 DB 739 ACATCCAAATCTGTTCTGGAGTCCCTGTCTGCTTCACTGAGTGGAGTCTGGGACCTCT 798  
 QY 601 TACTCTCTCAATATGACAGATGAGGCTGAAGATGCTGCCATTAATCTGACAGCAG 660  
 DB 799 TACTCTCTCAATATGACAGATGAGGCTGAAGATGCTGCCATTAATCTGACAGCAG 858

QY 661 TGGAGTAGTAACCACTCACTGCTGGTGGTGGAGCCAGACTGAGTGAAGCGGCGGCC 720  
 DB 859 TGGAGTAGTAATCACTCACTTGGTGGTGGAGCCAGACTGAGTGAAGCGGCGGCC 918  
 QY 721 GCAGAACAAAATCTATCTCAGAGAGGATCTGAATGGGGCGCTCCAGCAC 772  
 DB 919 GCAGAACAAAATCTATCTCAGAGAGGATCTGAATGGGGCGCTCCAGTATAC 970

# RESULT 5

AAE77806  
 ID AAE77806 standard; DNA; 1256 BP.

AAE77806;  
 AC

29-MAY-2001 (first entry)  
 DT

Fragment of pUBS520-ScFvOX coding for Pe1B-ScFvOXazolon fusion protein.  
 DE

Molecular chaperrone; Pe1B signal sequence; scFvOXazolon; ds.  
 KM

Unidentified.  
 OS

EP1077262-A1.  
 PN

21-FEB-2001.  
 PD

24-UTL-2000; 2000EP-00115839.  
 PF

29-UTL-1999; 99EP-00114811.  
 PR

(HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 PA

Ambrosius D, Rudolph R, Schaeffner J, Schwarz E;  
 PI

WPI; 2001-246712/26.  
 DR

P-PSDB; AAB74199.  
 DD

Producing naturally folded eukaryotic proteins e.g. antibodies,  
 PT interferon, hormones or proteases that contain two or several cysteines  
 PT linked by disulfide bridges comprises co-expression of a molecular  
 PT chaperrone.

Example 6; Page 17-18; 35pp; English.  
 PS

The present invention relates to a method for production of a naturally  
 CC folded eukaryotic protein containing two or more cysteines linked by  
 CC disulfide bridges. The method comprises co-expression and secretion into  
 CC the periplasm of a molecular chaperrone via an expression vector coding  
 CC for the chaperrone. The expression vector also encodes a signal sequence.  
 CC The method is useful for producing a naturally folded eukaryotic protein  
 CC such as an antibody, antibody fragment, interferon, protein hormone or a  
 CC protease containing two or several cysteines linked by disulfide bridges.  
 CC The present sequence is a fragment of expression vector pUBS520-ScFvOX,  
 CC which codes for a fusion protein composed of the Pe1B signal sequence and  
 CC ScFvOXazolon. This sequence was used in the method of the present  
 CC invention

SQ Sequence 1256 BP; 304 A; 321 C; 339 G; 292 T; 0 U; 0 Other;

Query Match 78.8%; Score 749.6; DB 4; Length 1256;

Best Local Similarity 98.2%; Pred. No. 4.8e-189;

Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGGCCGAGGTCAAGTGCAGAGAGTCAAGGAGGCTTAGTGAGGCTTGAGGCTCCCG 60  
 DB 199 ATGGCCGAGGTCAAGTGCAGAGAGTCTGGGGAGGCTTAGTGAGGCTTGAGGCTCCCG 258  
 QY 61 AAACCTCTCTGGGACCTCTGGATTCACCTTCAGTACGCTTGGAATGACCTGGGTGCT 120  
 DB 259 AAACCTCTCTGGGACCTCTGGATTCACCTTCAGTACGCTTGGAATGACCTGGGTGCT 318  
 QY 121 CAGGCTCCAGAGAGGGGCTGGAGTGGTGCATATATTAGTAGGAGTACATC 180



```
Db 319 CAGGCTCCAGAAAGGGCTGAGTGGGTGCGCATATTAGTAGGGCGATGTCACATC 378
Oy 181 TACTATGACAGACAGATGAAAGGAGCGATTACCATCTCCAGAGACAACTCCAAAGAACCC 240
Db 379 TACTATGACAGACAGATGAAAGGAGCGATTACCATCTCCAGAGACAACTCCAAAGAACCC 438
Oy 241 CTGTTCTTCGCAAAATGACAGCTTAAAGTCTGAGAGACAGCGTCATGTATTACTGTGCAGAA 300
Db 439 CTGTTCTTCGCAAAATGACAGCTTAAAGTCTGAGAGACAGCGTCATGTATTACTGTGCAGAA 498
Oy 301 GATTACGGGGCTTATTGGGGCCCAAGGACCAAGGTCAACCGTCTCCAGGTGGAAGCGGC 360
Db 499 GATTACGGGGCTTATTGGGGCCCAAGGACCAAGGTCAACCGTCTCCAGGTGGAAGCGGT 558
Oy 361 TCAGGCGGAGGTGGCTCTGCGCGGTGGCGGATGAGACATTGAGCTCAACCGATCTCCAGCA 420
Db 559 TCAGGCGGAGGTGGCTCTGCGCGGTGGCGGATGAGACATTGAGCTCAACCGATCTCCAGCA 618
Oy 421 ATCATGTCTGCATCTTCAGAGGAGAGAGGTCAACATGACCTGACAGTCCAGTTCAAGTGA 480
Db 619 ATCATGTCTGCATCTTCAGAGGAGAGAGGTCAACATGACCTGACAGTCCAGTTCAAGTGA 678
Oy 481 AGGTACATGAACCTGGTTCCAAAGAGTCAAGGCACTCCCCCAAAAGATGATTTATGAC 540
Db 679 AGGTACATGAACCTGGTTCCAAAGAGTCAAGGCACTCCCCCAAAAGATGATTTATGAC 738
Oy 541 ACATCCAAACTGTTCTTGGAGTCCCTGCTCCGCTTCAGTGGAGTGGGTCCGAGCTCT 600
Db 739 ACATCCAAACTGTTCTTGGAGTCCCTGCTCCGCTTCAGTGGAGTGGGTCCGAGCTCT 798
Oy 601 TACTCTTCACATACAGACAGCATGAGAGGTGAAGTCTGACACTTACTTCCAGCAG 660
Db 799 TACTCTTCACATACAGACAGCATGAGAGGTGAAGTCTGACACTTACTTCCAGCAG 858
Oy 661 TGGAGTAGTAACCACTCAACGTTGGTGTGCTGGAGCCAAAGCTGAGAGTGAAGGGCGGCG 720
Db 859 TGGAGTAGTAACCACTCAACGTTGGTGTGCTGGAGCCAAAGCTGAGAGTGAAGGGCGGCG 918
Oy 721 GCAGAACAAAATCTCATCTCAGAGAGGATCTGAATGGGGCCGTCGACGAGAC 772
Db 919 GCAGAACAAAATCTCATCTCAGAGAGGATCTGAATGGGGCCGTCATGTAAC 970
RESULT 6
AAF61193
ID AAF61193 standard; DNA; 1256 BP.
XX
AC AAF61193;
XX
DT 18-MAY-2001 (first entry)
XX
DE Expression plasmid pUBS520-ScFvOx DNA.
XX
KM Chaperone protein; periplasm; antibody production; protein production;
interferon production; protease production; ds.
XX
OS Escherichia coli.
XX
XX Synthetic.
XX
PH Key Location/Qualifiers
FT CDS 199..966
FT /tag= a
FT /product= "pUBS520-ScFvOx"
XX
PN EPI077263-A1.
PD 21-FEB-2001.
XX
PF 29-JUL-1999; 99EP-00114811.
XX
PR 29-JUL-1999; 99EP-00114811.
XX
```

```
PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
DR WPI: 2001-204356/21.
DR P-PSDB; AAB70769.
XX
PT Preparation of naturally folded eukaryotic proteins, e.g. antibodies, by
PT simultaneous expression of a chaperone protein, allows simple recovery
PT from periplasm or medium.
XX
XX Example 6; Page 18-20; 36pp; German.
CC This invention describes a novel method for preparing a naturally folded
CC eukaryotic polypeptide (I) that contains two or more disulfide-bridged
CC Cys residues by culturing prokaryotic cells that contain an expression
CC vector for (I) including a prokaryotic signal sequence at its N-terminus
CC and a nucleic acid (II) that secretes a chaperone protein (III) into the
CC periplasm. (I) is secreted into the periplasm or medium; the signal
CC peptide is then cleaved and (I) isolated from the periplasm or medium.
CC The method is used for production of antibody, interferon, protein
CC hormone or protease. Expression of (III) increases the yield of (I). The
CC method is simple and eliminates time-consuming in vitro processing
CC operations such as dissolution of inclusion bodies, reduction and
CC refolding. (III) protects (I) against agglomeration and promotes their
CC natural conformation
XX
SQ Sequence 1256 BP; 304 A; 320 C; 339 G; 293 T; 0 U; 0 Other;
Query Match 78.8%; Score 749.6; DB 4; Length 1256;
Best Local Similarity 98.2%; Pred. No. 4,8e-189;
Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Oy 1 ATGCGCGAGGTCAAGCTCAGAGAGTCAAGGGAGGCTTATGTCACCTCGAGAGGTCCTCGG 60
Db 199 ATGCGCGAGGTCAAGCTCAGAGAGTCAAGGGAGGCTTATGTCACCTCGAGAGGTCCTCGG 258
Oy 61 AAACCTCTCTGCGAGCCTCTGAGATTCACTTCACTGTAAGTGAATGCACTGGGTTGCT 120
Db 259 AAACCTCTCTGCGAGCCTCTGAGATTCACTTCACTGTAAGTGAATGCACTGGGTTGCT 318
Oy 121 CAGGCTCAGAGAAAGGGCTGAGTGGGTGCGCATTTATAGTAGGAGTAGTACATC 180
Db 319 CAGGCTCAGAGAAAGGGCTGAGTGGGTGCGCATTTATAGTAGGAGTAGTACATC 378
Oy 181 TACTATGACAGACAGATGAAAGGAGCGATTCAACATCTCCAGAGCAATCCCAAGAACCC 240
Db 379 TACTATGACAGACAGATGAAAGGAGCGATTCAACATCTCCAGAGCAATCCCAAGAACCC 438
Oy 241 CTGTTCTTCGCAAAATGACAGCTTAAAGTCTGAGAGACAGCGTCATGTATTACTGTGCAGAA 300
Db 439 CTGTTCTTCGCAAAATGACAGCTTAAAGTCTGAGAGACAGCGTCATGTATTACTGTGCAGAA 498
Oy 301 GATTACGGGGCTTATTGGGGCCCAAGGACCAAGGTCAACCGTCTCCAGGTGGAAGCGGC 360
Db 499 GATTACGGGGCTTATTGGGGCCCAAGGACCAAGGTCAACCGTCTCCAGGTGGAAGCGGT 558
Oy 361 TCAGGCGGAGGTGGCTCTGCGCGGTGGCGGATGAGACATTGAGCTCAACCGATCTCCAGCA 420
Db 559 TCAGGCGGAGGTGGCTCTGCGCGGTGGCGGATGAGACATTGAGCTCAACCGATCTCCAGCA 618
Oy 421 ATCATGTCTGCATCTTCAGAGGAGAGAGGTCAACATGACCTGACAGTCCAGTTCAAGTGA 480
Db 619 ATCATGTCTGCATCTTCAGAGGAGAGAGGTCAACATGACCTGACAGTCCAGTTCAAGTGA 678
Oy 481 AGGTACATGAACCTGGTTCCAAAGAGTCAAGGCACTCCCCCAAAAGATGATTTATGAC 540
Db 679 AGGTACATGAACCTGGTTCCAAAGAGTCAAGGCACTCCCCCAAAAGATGATTTATGAC 738
Oy 541 ACATCCAAACTGTTCTTGGAGTCCCTGCTCCGCTTCAGTGGAGTGGGTCTGGAACCTCT 600
Db 739 ACATCCAAACTGTTCTTGGAGTCCCTGCTCCGCTTCAGTGGAGTGGGTCTGGAACCTCT 798
Oy 601 TACTCTTCACATACAGACAGCATGAGAGGTGAAGTCTGCACTTACTTCCAGCAG 660
```



Db 799 TACTCTTCACATCATGACGATGAGGCTGAAGATGCTGCCATTATTACTGCCAGCAG 858  
 Qy 661 TGGAGTAGTAACCCACTACGCTTCGGTGTGGAGCCAAAGCTGGAGCTGAACCGGGCGGCC 720  
 Db 859 TGGAGTAGTAACCCACTACGCTTCGGTGTGGAGCCAAAGCTGGAGCTGAACCGGGCGGCC 918  
 Qy 721 GCAGAACAAAACTCATCTCAGAAAGAGATCTGAATGGGGCCGTCGACGAAC 772  
 Db 919 GCAGAACAAAACTCATCTCAGAAAGAGATCTGAATGGGGCCGTCATATGTAAC 970  
 RESULT 7  
 AAD02212  
 ID AAD02212 standard; DNA; 1256 BP.  
 AC AAD02212;  
 XX 28-MAR-2001 (first entry)  
 XX Partial PUBS520-pin-scFvOX encoding PelB-scFvOXazolon fusion protein.  
 XX Secreted protein; plasmid PUBS520-pin-scFvOX; chaparone; interferon;  
 KW protease; hormone; naturally folded protein; fusion protein; DnaJ;  
 KW heat shock protein; HSP; pectate lyase B; PelB; hapten; lac promoter;  
 KW single-chain Fv-fragment Oxazolone; scFvOXazolon;  
 KW thyroid stimulating hormone; TSH; ds.  
 XX Pectobacterium carotovorum.  
 OS Escherichia coli.  
 OS unidentified.  
 OS Chimeric.  
 FH Key Location/Qualifiers  
 FH CDS 199..966  
 FT /tag= a  
 FT /product= "Erwinia carotovora PelB signal sequence-  
 FT single-chain Fv-fragment Oxazolone fusion protein"  
 FT /note= "CDS has an additional stop codon at position 967-  
 FT 969"  
 XX BP1054063-A2.  
 XX 22-NOV-2000.  
 PD 19-APR-2000; 2000BP-00108505.  
 PF 26-APR-1999; 99BP-00107412.  
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.  
 PA Ambrosius D, Rudolph R, Schaeffner J, Schwarz E;  
 PI WPI; 2001-033777/05.  
 DR P-PSDB; AA972020.  
 XX Producing water-soluble, naturally folded, and secreted eukaryotic  
 PT polypeptide, involves culturing prokaryotic cells containing an  
 PT expression vector encoding the polypeptide in the presence of arginine or  
 PT a specific compound.  
 XX Example 6; Page 20-22; 35p; English.  
 XX The patent discloses a method for the production of a water-soluble,  
 CC naturally folded and secreted eukaryotic proteins in prokaryotic cells.  
 CC The method involves culturing the prokaryotic cells, containing an  
 CC expression vector encoding the desired protein and the prokaryotic signal  
 CC sequence, in the presence of an additive, e.g., L-arginine. The signal  
 CC sequence mediates the secretion of the desired protein into the  
 CC periplasm, where folding of the protein takes place. The prokaryotic cell  
 CC also contains an expression vector encoding a molecular chaperone, e.g.,  
 CC DnaJ and heat shock protein 25 (HSP25). The simultaneous co-  
 CC overexpression and co-secretion of molecular chaperones in the periplasm  
 CC improves the yield of functionally folded protein. The above method

CC recombinantly produces a high yield of eukaryotic secreted proteins in  
 CC prokaryotes. The method is useful for producing eukaryotic proteins such  
 CC as an antibody, antibody fragment, interferon, protein hormone or a  
 CC protease. The present sequence is a part of an expression plasmid.  
 CC PUBS520-pin-scFvOX, which comprises the lac promoter, the Erwinia  
 CC carotovora pectate lyase B (PelB) signal sequence and the scfvox gene.  
 CC The single-chain Fv-fragment, which is directed against the hapten  
 CC oxazolone (scFvOXazolon), is an antibody fragment against thyroid  
 CC stimulating hormone (TSH). The co-expression of scFvOX which has no  
 CC chaperone properties is used as a negative control  
 XX Sequence 1256 BP; 304 A; 320 C; 339 G; 293 T; 0 U; 0 Other;  
 SQ  
 Query Match 78.8%; Score 749.6; DB 5; Length 1256;  
 Best Local Similarity 98.2%; Pred. No. 4.8e-189;  
 Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 Qy 1 ATGGCCGAGGTCAAGCTGCAAGAGTCAAGGGGAGCTTATGTCAGCTGGAGGGTCCCG 60  
 Db 199 ATGGCCGAGGTCAAGCTGCAAGAGTCTGGGGGAGCTTATGTCAGCTGGAGGGTCCCG 258  
 Qy 61 AAATCTCTGTGACGCTCTGATTCACCTTCAGTAGCTTTGGATATGACTGGGTGCT 120  
 Db 259 AAATCTCTGTGACGCTCTGATTCACCTTCAGTAGCTTTGGATATGACTGGGTGCT 318  
 Qy 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCATATATTAGTAGTGGAGTAGTACATC 180  
 Db 319 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCATATATTAGTAGTGGAGTAGTACATC 378  
 Qy 181 TACTATGACAGACAGTGAAGGAGCATTACCATCTCCAGAGCAATCCCAAGAACCC 240  
 Db 379 TACTATGACAGACAGTGAAGGAGCATTACCATCTCCAGAGCAATCCCAAGAACCC 438  
 Qy 241 CTGTTCTCTGCAATGACAGTCTTAAGTCTGAGAGACAGGTATTAAGTGAAGA 300  
 Db 439 CTGTTCTCTGCAATGACAGTCTTAAGTCTGAGAGACAGGTATTAAGTGAAGA 498  
 Qy 301 GATTACGGGGCTTAATGGGGCCAAAGGAGCAACAGTCAAGCTCTCTCAGGTGAGGCGGC 360  
 Db 499 GATTACGGGGCTTAATGGGGCCAAAGGAGCAACAGTCAAGCTCTCTCAGGTGAGGCGGC 558  
 Qy 361 TCAGGGGAGAGTGGCTCTGGCGGTGGCGAGATGGAGCATTTAGAGTCAACCATGCTCCAGCA 420  
 Db 559 TCAGGGGAGAGTGGCTCTGGCGGTGGCGAGATGGAGCATTTAGAGTCAACCATGCTCCAGCA 618  
 Qy 421 ATCATGTCTGACATCTCCAGAGGAGAGGATCAACATGACCTGACAGTCCAGTCAAGTGA 480  
 Db 619 ATCATGTCTGACATCTCCAGAGGAGAGGATCAACATGACCTGACAGTCCAGTCAAGTGA 678  
 Qy 481 AGGTACATGAATCTGTTCCAAAGAGTCAAGGACCTCCCAAAAGATGATTTATGAC 540  
 Db 679 AGGTACATGAATCTGTTCCAAAGAGTCAAGGACCTCCCAAAAGATGATTTATGAC 738  
 Qy 541 ACATCCAAATCTGTTCTGGAGTCCCTGCTGCTTCAATGGACAGTGGAGCTCT 600  
 Db 739 ACATCCAAATCTGTTCTGGAGTCCCTGCTGCTTCAATGGAGTGGAGCTCT 798  
 Qy 601 TACTCTTCACATCATGACGATGAGGCTGAAGATGCTGCCATTAATCTCCACAGC 660  
 Db 799 TACTCTTCACATCATGACGATGAGGCTGAAGATGCTGCCATTAATCTCCACAGC 858  
 Qy 661 TGGAGTAGTAACCCACTACGCTTCGGTGTGGAGCCAAAGCTGGAGCTGAACCGGGCGGCC 720  
 Db 859 TGGAGTAGTAACCCACTACGCTTCGGTGTGGAGCCAAAGCTGGAGCTGAACCGGGCGGCC 918  
 Qy 721 GCAGAACAAAACTCATCTCAGAAAGAGATCTGAATGGGGCCGTCGACGAAC 772  
 Db 919 GCAGAACAAAACTCATCTCAGAAAGAGATCTGAATGGGGCCGTCATATGTAAC 970  
 RESULT 8  
 ACC84876  
 ID ACC84876 standard; DNA; 843 BP.

[illegible]

| QY       | 244 | TTCTGCAATATGACCAAGTCTTAAGTCTGAGACACGGTCAATGATTATTAATCTGTGCAGAGAT | 307 |
|----------|-----|--|-----|
| Db       | 310 | TTCTGCAAAATGACCAAGTCTTAAGTCTGAGACACGGCATGATTATTAATCTGTGCAGAGAT   | 368 |
| QY       | 304 | TACGGGGCTTATTGGGGGCGCAAGGACCAAGGATCACGGTCTCCCTCAGGTGGAGGCGGCTCA  | 363 |
| Db       | 370 | TACGGGGCTTATTGGGGGCGCAAGGACCAAGGATCACGGTCTCCCTCAGGTGGAGGCGGCTCA  | 422 |
| QY       | 364 | GCGGAGGTGCGTCTGTGGCGGTGGCGGATCGGACATTGAGCTCACCCAGTCTCCAGCAATC    | 423 |
| Db       | 430 | GCGGAGGTGCGTCTGTGGCGGTGGCGGATCGGACATTGAGCTCACCCAGTCTCCAGCAATC    | 489 |
| QY       | 424 | ATGTCTGATCTTCCAGGGGAGAGGGTACACATGATCCTGGAGTCCAGTTCAATGTAAAG      | 483 |
| Db       | 430 | ATGTCTGATCTTCCAGGGGAGAGGGTACACATGATCCTGGAGTCCAGTTCAATGTAAAG      | 549 |
| QY       | 484 | TACATGAAGTGGTTTCCAAAGATCAGGACCTCCCCCAAAAGATGATTTATGACACA         | 543 |
| Db       | 550 | TACATGAAGTGGTTTCCAAAGATCAGGACCTCCCCCAAAAGATGATTTATGACACA         | 609 |
| QY       | 544 | TCGAACTGTCTTCTGAGTCCCTGCTGCTTCAGTGAGTGGGTCTGGGACCTCTTAC          | 603 |
| Db       | 610 | TCGAACTGTCTTCTGAGTCCCTGCTGCTTCAGTGAGTGGGTCTGGGACCTCTTAC          | 665 |
| QY       | 604 | TCTCTCAGAAATGAGAGCATGGAGGCTGAAGATGCTGCCACTTAATACGCCAGCAGTGG      | 665 |
| Db       | 670 | TCTCTCAGAAATGAGAGCATGGAGGCTGAAGATGCTGCCACTTAATACGCCAGCAGTGG      | 729 |
| QY       | 664 | AGTAGTAACCCACATCAGTTCGGTCTGGGACCAAGCTGGAGCTGAAGCGGGCGCGCA        | 722 |
| Db       | 730 | AGTAGTAATTCACATCACTTTGGGTCTGGGACCAAGCTGGAGCTGAAGCGGGCGCGCA       | 789 |
| QY       | 724 | GAAACAAAATCTATCTCAGAAAGGATCTGAATGGGCGCTGACGAAAC                  | 772 |
| Db       | 730 | GAAACAAAATCTATCTCAGAAAGGATCTGAATGGGCGCTGACGAAAC                  | 838 |
| RESULT 9 |     |  |     |
| AAK98639 | ID  | AAK98639 standard; DNA, 726 BP.                                  |     |
| XX       | AC  | AAK98639;  |     |
| XX       | DT  | 03-MAY-2002 (first entry)  |     |
| DE       | XX  | scFv antibody coding sequence.                                   |     |
| XX       | XX  | Ketone binding protein; oxazole; pathogen resistance; virucide;  |     |
| KW       | XX  | fungicide; antibacterial; scFv antibody; gene; ds.               |     |
| XX       | OS  | Unidentified.  |     |
| FX       | Key | Location/Qualifiers  |     |
| FT       | CDS | 1..726   |     |
| FT       |     | /*tag= a   |     |
| FT       |     | /product= "scFv antibody"  |     |
| FT       |     | /transl_except= (pos:409..411,aa:Gly)                            |     |
| FT       |     | /transl_except= (pos:601..603,aa:Val)                            |     |
| PN       | XX  | WO200204020-A2.  |     |
| PD       | XX  | 17-JAN-2002.   |     |
| PF       | XX  | 17-MAY-2001; 2001WO-DE001916.                                    |     |
| PR       | XX  | 12-JUL-2000; 2000DE-01033750.                                    |     |
| PA       | XX  | (MPBC-) MPB COLOGNE GMBH.  |     |
| PI       | XX  | Duering K, Brinkmann O;  |     |
| DR       | XX  | WPI; 2002-154868/20.   |     |

DR P-PSDB, AAM48925.

XX Imparting pathogen resistance to plants and animals, comprises using a  
PT polypeptide that binds to an oxazole-derived ketone, optionally expressed  
PT from nucleic acid.

XX Example 3, Fig 1, 20pp; German.

XX The present invention relates to the use of a polypeptide that binds a  
CC ketone derived from oxazole for generating pathogen resistance in an  
CC organism. This can be used to impart resistance to pathogens (bacteria,  
CC fungi or viruses) to a wide range of plants (e.g. cereals, sugar beet,  
CC tobacco etc.), humans, farm animals and pets. Exemplified are activities  
CC against Phytophthora infestans and Erwinia carotovora in potatoes and  
CC Staphylococcus aureus in mice. The present sequence is a sc-Fv antibody  
CC coding sequence, which encodes a protein capable of binding 2-  
CC phenylloxazol-5-one

XX Sequence 726 BP; 166 A; 184 C; 214 G; 162 T; 0 U; 0 Other;

XX Query Match 73.5%; Score 699.4; DB 6; Length 726;

XX Best Local Similarity 97.8%; Pred. No. 9e-176;  
Matches 709; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCCGAGGTCAAGCTGACAGAGTCAAGGAGGCTTATGTCAGCCTGAGGGTCCCG 60  
DB 1 ATGGCTGATGTGACGCTGGTGAAGTCTGGGGAGGCTTATGTCAGCCTGAGGGTCCCG 60  
QY 61 AAACCTCTCTGTGACAGCCTCTGGAATTCATTCAAGTCTTGAATGACCTGGGTTCTG 120  
DB 61 AAACCTCTCTGTGACAGCCTCTGGAATTCATTCAAGTCTTGAATGACCTGGGTTCTG 120  
QY 121 CAGGCTCCAGAGAAAGGGCTGAGTGGGTGCGATATATTAGTAGGAGTAGTACCATC 180  
DB 121 CAGGCTCCAGAGAAAGGGCTGAGTGGGTGCGATATATTAGTAGGAGTAGTACCATC 180  
QY 181 TACTATGACAGACAGATGAAGGAGCATTCACATCTCCAGAGACATCCAGAACACC 240  
DB 181 TACTATGACAGACAGATGAAGGAGCATTCACATCTCCAGAGACATCCAGAACACC 240  
QY 241 CTGTTCTCTGCAATGACCAAGTCTAAGTCTGAGAGACAGGTATATTAAGTGTGACAA 300  
DB 241 CTGTTCTCTGCAATGACCAAGTCTAAGTCTGAGAGACAGGTATATTAAGTGTGACAA 300  
QY 301 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCACCGTCTCTCAGGTGAGGGCGC 360  
DB 301 GATTACGGGGCTTATTTGGGGCCAAAGGACCAAGGTCACCGTCTCTCAGGTGAGGGCGC 360  
QY 361 TCAGGGGAGGTGGCTCTGGCCGGTGGCGATTCGAGATTCGACCCAGTCTCCAGCA 420  
DB 361 TCAGGGGAGGTGGCTCTGGCCGGTGGCGATTCGAGATTCGACCCAGTCTCCAGCA 420  
QY 421 ATCATATCTGTCATCTCCAGGGGAGAGGGTCAACATGACCTGCGAGTCCAGTTAAAGTGA 480  
DB 421 ATCATATCTGTCATCTCCAGGGGAGAGGGTCAACATGACCTGCGAGTCCAGTTAAAGTGA 480  
QY 481 AGGTACATGAACTGTTCCAAAGAGTCAAGGACCTCCCGCCAAAGATGAGATTTATGAC 540  
DB 481 AGGTACATGAACTGTTCCAAAGAGTCAAGGACCTCCCGCCAAAGATGAGATTTATGAC 540  
QY 541 ACATCCAAACCTGTTCTGGAGTCCCTGCTGCTTCAAGTGGAGTGGGAGCTCTT 600  
DB 541 ACATCCAAACCTGTTCTGGAGTCCCTGCTGCTTCAAGTGGAGTGGGAGCTCTT 600  
QY 601 TACTCTCTCAATCAGCAGATGAGAGGCTGAAGATGCTGCCACTTAATCTGCCAGAG 660  
DB 601 TACTCTCTCAATCAGCAGATGAGAGGCTGAAGATGCTGCCACTTAATCTGCCAGAG 660  
QY 661 TGGAGTAGTAACCATCTACGTTCCGGTGTGGAGCAAGCTGAGAGTGAAGCGGGGGCC 720  
DB 661 TGGAGTAGTAACCATCTACGTTCCGGTGTGGAGCAAGCTGAGAGTGAAGCGGGGGCC 720  
QY 721 GCAGA 725

DB 721 GCATA 725

RESULT 10

AA080468

ID AA080468 standard; DNA; 738 BP.

AC AA080468;

DT 25-MAR-2003 (revised)

DT 13-SEP-1993 (first entry)

DE Single chain antibody (scfv) which binds to phenylloxazolone.

KM Genetic selection; ligand binding protein; cholera toxin; promoter;

KM regulatory domain; db.

OS Synthetic.

FT Key Location/Qualifiers

FT CDS 1..723

FT /tag= a

FT /product= "single chain antibody."

DE4319296-A1.

15-DEC-1994.

10-JUN-1993; 93DE-04319296.

10-JUN-1993; 93DE-04319296.

(BEHM) BEHRINGER AG.

Fritz H, Hennecke F, Kolmar H,

WPI; 1995-023689/04.

P-PSDB; AAR68613.

Genetic selection of ligand binding proteins in microorganisms - by

extracytoplasmic protein presentation, then use of ligand binding to

express a detectable or selectable function.

Example 2.2; Fig 4; 25pp; German.

Genetic selection in microorganisms, for ligand binding proteins (LBP)

comprises: extracytoplasmic presentation of LBP and; using the signal of

ligand binding (by signal transduction on the biosynthetic machinery of

the microorganisms) to express a detectable or selectable function.

Microorganisms for this process include a genetically stable

detection/selection and are transformed with a replicon encoding a fusion

protein consisting of the LBP, a transmembrane helix and regulatory

domain. The detection/selection function is expression of a beta-

galactosidase gene, integrated into the chromosome and under the control

of the ctx (cholera toxin) promoter. The transmembrane helix is taken

from the toxR gene. Four primers (AA080457-60) were used in the

DB

721 GCATA 725

RESULT 10

AA080468

ID AA080468 standard; DNA; 738 BP.

AC AA080468;

DT 25-MAR-2003 (revised)

DT 13-SEP-1993 (first entry)

DE Single chain antibody (scfv) which binds to phenylloxazolone.

KM Genetic selection; ligand binding protein; cholera toxin; promoter;

KM regulatory domain; db.

OS Synthetic.

FT Key Location/Qualifiers

FT CDS 1..723

FT /tag= a

FT /product= "single chain antibody."

DE4319296-A1.

15-DEC-1994.

10-JUN-1993; 93DE-04319296.

10-JUN-1993; 93DE-04319296.

(BEHM) BEHRINGER AG.

Fritz H, Hennecke F, Kolmar H,

WPI; 1995-023689/04.

P-PSDB; AAR68613.

Genetic selection of ligand binding proteins in microorganisms - by

extracytoplasmic protein presentation, then use of ligand binding to

express a detectable or selectable function.

Example 2.2; Fig 4; 25pp; German.

Genetic selection in microorganisms, for ligand binding proteins (LBP)

comprises: extracytoplasmic presentation of LBP and; using the signal of

ligand binding (by signal transduction on the biosynthetic machinery of

the microorganisms) to express a detectable or selectable function.

Microorganisms for this process include a genetically stable

detection/selection and are transformed with a replicon encoding a fusion

protein consisting of the LBP, a transmembrane helix and regulatory

domain. The detection/selection function is expression of a beta-

galactosidase gene, integrated into the chromosome and under the control

of the ctx (cholera toxin) promoter. The transmembrane helix is taken

from the toxR gene. Four primers (AA080457-60) were used in the

Sequence 738 BP; 170 A; 190 C; 209 G; 169 T; 0 U; 0 Other;

Query Match 68.9%; Score 655.4; DB 2; Length 738;

Best Local Similarity 95.0%; Pred. No. 4.6e-164;  
Matches 677; Conservative 0; Mismatches 36; Indels 0; Gaps 0;



```

Db      1 GATGTGACAGCTGTGGAGTCTGTGGGGAGGCTTATGTGACCTGTGGAGGCTCCGGAAATCTC 60
Oy      67 TCCTGTGACAGCTCTGGATTCACTTCACTAGAGCTTGAATGACATGGGTCGTACAGCT 126
Db      61 TCCTGTGACAGCTCTGGATTCACTTCACTAGAGCTTGAATGACATGGGTCGTACAGCT 120
Oy      127 CCAGAGAAAGGGGTGGAGTGGGTGCGATATATTAGTAGTGGAGTAGTACCATCTAT 186
Db      121 CCAGAGAAAGGGGTGGAGTGGGTGCGATATATTAGTAGTGGAGTAGTACCATCTAT 180
Oy      187 GCAGACACAGTGAAGGACGATTCACCATCTCCAGAGACATCCCAAGAACACCTCTTTC 246
Db      181 GCAGACACAGTGAAGGACGATTCACCATCTCCAGAGACATCCCAAGAACACCTCTTTC 240
Oy      247 CTGCAAAATGACAGTCTTAAGTCTGAGGACACGGATCATGTATTAATCTGTGCAAGATTC 306
Db      241 CTGCAAAATGACAGTCTTAAGTCTGAGGACACGGATCATGTATTAATCTGTGCAAGATTC 300
Oy      307 GGGGCTTATTTGGGGCCAGAGGACACAGCTCACCGTCTCTCAGTGGAGGGCGCTCAGGC 366
Db      301 GGGGCTTATTTGGGGCCAGAGGACCTGTGTCAGTCTGTGACAGTGGTGGTGGTACGGGT 360
Oy      367 GGAGGTGCTCTGCGCGTGGCGGATCGACATTTGAGCTCACCCAGTCTCCAGCAATCATG 426
Db      361 GGAGGTGCTCTGCGCGTGGCGGATCGACATTTGAGCTCACCCAGTCTCCAGCAATCATG 420
Oy      427 TCTGATCTCCAGGGGAGAGGCTCACCATGACCTGACAGTCCAGTCAAGTAAAGTAC 486
Db      421 TCTGATCTCCAGGGGAGAGGCTCACCATGACCTGACAGTCCAGTCAAGTAAAGTAC 480
Oy      487 ATGAACTGTGTTCCAAAGAGTCAAGGACCTCCCAAAAGATGATTTATGACATCC 546
Db      481 ATGAACTGTGTTCCAAAGAGTCAAGGACCTCCCAAAAGATGATTTATGACATCC 540
Oy      547 AAACGTCTCTTCTGAGATCCCTGCTGCTTCACTGAGGAGTGGGTCTGGACCTTTACTCT 606
Db      541 AAACGTCTCTTCTGAGATCCCTGCTGCTTCACTGAGGAGTGGGTCTGGACCTTTACTCT 600
Oy      607 CTCACAAATGACGATGAGGAGGCTGAAGATGCTGCCACTTACTGACGACGATGAGT 666
Db      601 CTCACAAATGACGATGAGGAGGCTGAAGATGCTGCCACTTACTGACGACGATGAGT 660
Oy      667 AGTAAACCACTCACGTTCTGGTCTGGAGCCAAAGCTGAGAGTAAAG 713
Db      661 AGTAAATCCACTCACGTTCTGGTCTGGAGCCAAAGCTGAGAGTAAAG 707

RESULT 12
AAQ37460
ID      AAQ37460 standard; DNA; 672 BP.
XX
AC      AAQ37460;
XX
XX      25-MAR-2003 (revised)
DT      19-JUN-1993 (first entry)
XX
DE      VH NQ10/12.5-Vk NQ10/12.5 linked coding sequence.
KW      primer; human; immunoglobulin; Ig; variable region; VH; VI; Ck; JH;
KW      lymphocyte; vector; soluble; antibody; phage; linker; back; VH3; nested;
KW      in-cell PCR; cloning; polymorphic; TCR V; antiphenyloxazalone; hybridoma;
KW      NQ2/12.4; NQ10/12.5; 86.
XX
OS      Synthetic.
XX
FH      Key Location/Qualifiers
FT      misc_RNA 1..342
FT      /*tag= a
FT      /label= VH_NQ10/12.5
FT      primer_bind 5..26
FT      /*tag= b
FT      /label= NQ10/12.5_BACK
FT      primer_bind 38..58

```

```

FT      /*tag= c
FT      /label= NQ10/12.5BKNES
FT      primer_bind complement(41..60)
FT      /*tag= d
FT      /label= NQHPRB
FT      primer_bind complement(310..355)
FT      /*tag= e
FT      /label= MovHlnk4
FT      primer_bind 330..373
FT      /*tag= h
FT      /label= MovKappaInk4
FT      misc_feature 343..348
FT      /*tag= f
FT      /note= "Linker Peptide"
FT      misc_RNA 349..672
FT      /*tag= g
FT      /label= Vk_NQ10/12.5
FT      primer_bind complement(438..456)
FT      /*tag= i
FT      /label= NQ10KPRB
FT      primer_bind complement(628..648)
FT      /*tag= j
FT      /label= MOKappa5FORNES
FT      primer_bind complement(650..672)
FT      /*tag= k
FT      /label= MOKappa5FOR2

MO9303151-A1.
XX
PD      18-FEB-1993.
XX
PF      10-AUG-1992; 92WC-GB001483.
XX
PR      10-AUG-1991; 91GB-00017352.
PR      11-JUN-1992; 92GB-00012419.
XX
PA      (MED1-) MEDICAL RES COUNCIL.
XX
PI      Embleton MJ, Gorochov G, Jones PT, Winter GP;
XX      WPI, 1993-076508/09.
DR      P-PSDB; AAR32841.
XX
XX      Treatment of cell populations, partic. hybridomas - to link together
XX      copies of 2 or more non-contiguous DNA sequences to facilitate analysis.
XX
XX      Disclosure: Fig 3; 72pp; English.
XX
XX      The sequences given in AAQ37459-62 show the nucleotide sequences of
XX      mature heavy chain VH domains and the Vk light chain genes of the
XX      antiphenyloxazalone hybridomas NQ2/12.4 and NQ10/12.5 which have been
XX      linked via a linker peptide by in-cell PCR. The cDNA was synthesised
XX      using forward primers annealing to the Ck gene and the VH segment,
XX      followed by assembly with linker primers, VH back primers based on the
XX      VH3 leader sequence and a forward Ck primer nested in respect to the
XX      primer used for cDNA. The assembled product within the cells is then
XX      amplified with nested primers annealing to the 5' end of the VH gene and
XX      the 3' end of the Jk segment. In-cell PCR may be used to determine gene
XX      linkage analysis, particularly for the cloning of gene combinations that
XX      are polymorphic within a population of cells, such as the rearranged
XX      Ck genes for Ig or TCR V regions. (Updated on 25-MAR-2003 to correct PN
XX      field.)
XX
SO      Sequence 672 BP; 161 A; 169 C; 184 G; 158 T; 0 U; 0 Other;

Query Match 61.1%; Score 580.6; DB 2; Length 672;
Best Local Similarity 91.1%; Pred. No. 3.6e-144;
Matches 644; Conservative 0; Mismatches 24; Indels 39; Gaps 1;

Oy      7 GAGGTAACTGACAGAGTCAAGGGGAGGCTTATGTGACCTGTGAGGCTCCGGAAATCTC 66
Db      1 GATGTGACAGCTGTGGAGTCTGTGGGGAGGCTTATGTGACCTGTGGAGGCTCCGGAAATCTC 60

```

```

QY 67 TCCTGTGACGCTCTGAGATTCACTTTCAGTTCGATGCACTGGGTTGTCAGGCT 126
DB 61 TCCTGTGACGCTCTGAGATTCACTTTCAGTTCGATGCACTGGGTTGTCAGGCT 120
QY 127 CCAGAGAAAGGGGCTGGAGTGGGTCGCACTATTATTAGTAGGGCAGTAGTACCATCTACTAT 186
DB 121 CCAGAGAAAGGGGCTGGAGTGGGTCGCACTATTATTAGTAGGGCAGTAGTACCATCTACTAT 180
QY 187 GCAGACACAGTAGAAGGGACGATTCACATCTCCAGAGACATCCCAAGAACACCCCTGTTTC 246
DB 181 GCAGACACAGTAGAAGGGACGATTCACATCTCCAGAGACATCCCAAGAACACCCCTGTTTC 240
QY 247 CTGCAAAATACCAAGTCTTAAGTCTTAGAGACAGGTCATGATTTACTGTGCAAGAGATTAC 306
DB 241 CTGCAAAATACCAAGTCTTAAGTCTTAGAGACAGGTCATGATTTACTGTGCAAGAGATTAC 300
QY 307 GGGGGCTATTGGGGGCCAAGGAGCAACGGTCACCGTCTCCTCAGGTGGAGGCGGCTCAGGC 366
DB 301 GGGGGCTATTGGGGGCCAAGGAGCTGAGTCACCTGTCTGCAAGTAC----- 347
QY 367 GGAGGTGCTCTGGCGGTGGCGGATCGGACATTTAGAGCTCACCAGTCTCCAGCAATCATG 426
DB 348 -----CAAATTGTTCTCACCCAGTCTCCAGCAATCATG 381
QY 427 TCTGCATCTCCAGGGAGAGGGTCACTCATGACCTGCAAGTCCAGATTCAAGTGAAGTAC 486
DB 382 TCTGCATCTCCAGGGAGAGGGTCACTCATGACCTGCAAGTCCAGATTCAAGTGAAGTAC 441
QY 487 ATGAAGTGGTTCACACAGAAAGTCAGGCACTCCCCCAAAAGATGATTATAGACATATCC 546
DB 442 ATGAAGTGGTTCACACAGAAAGTCAGGCACTCCCCCAAAAGATGATTATAGACATATCC 501
QY 547 AAAGTGTCTTCTGAGAGTCCCTGCTGCTTCAGTGGCAGTGGGTTCTGGGACCTTTACTCT 606
DB 502 AAAGTGTCTTCTGAGAGTCCCTGCTGCTTCAGTGGCAGTGGGTTCTGGGACCTTTACTCT 561
QY 607 CTCACAAATAGACAGATGAGAGCTGAAGATGCTGCCACTTACTACTGCGACAGATGAGAGT 666
DB 562 CTCACAAATAGACAGATGAGAGCTGAAGATGCTGCCACTTACTACTGCGACAGATGAGAGT 621
QY 667 AGTAAACCCATCAGCTTGGGTGCTGGGACCAAGCTGGAGCTGAAACG 713
DB 622 AGTAAATCCATCAGCTTGGGTGCTGGGACCAAGCTGGAGCTGAAACG 668

RESULT 13
ID AAX01651
AAX01651 strand: DNA; 1725 BP.
AAX01651;
10-MAY-1999 (first entry)
Bispecific tetraivalent antibody B1TAB724-1G10H6 DNA.
DE B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2; CD86;
KM T cell activation; inhibitor; graft versus host disease;
KM transplant rejection; allograft rejection; autoimmune disease; allergy;
KM therapy; human; bispecific tetraivalent antibody; B1TAB;
KM B1TAB724-1G10H6, ss.
XX
OS Mus sp.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1726
FT FT /*tag= a
FT FT /transl_except= (pos:780..782, aa:11e)
FT FT /transl_except= (pos:967..969, aa:Xaa)
FT FT /note= "Xaa= Gln-Leu-Val-Gln-Val-Gln"
FT sig_peptide 1..72

```

```

FT /*tag= b
FT /note= "pe1B signal sequence"
FT mat_peptide 73..1726
FT FT /*tag= c
FT PN MO9858965-A2.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98MO-EP003791.
XX
PR 20-JUN-1997; 97BP-00870092.
XX
PA (INNO-) INNOGENETICS NV.
PI Lorre K, Sablon E, Buysse M, Boesman A;
DR MPI. 1999-105615/09.
DR P-PSDB; AAM90217.
XX
PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PS immune diseases including allograft rejection.
XX Example 7.1; Fig 15; 182pp; English.
XX This DNA sequence encodes the bispecific tetraivalent antibody B1TAB724-
XX 1G10H6 (see AAM90217) comprising the anti-B7.1 murine monoclonal antibody
XX (MAB) B7-24 VH region joined, via a flexible synthetic linker, to the B7-
XX 24 VL region, the human IgG3 hinge region, a helix-turn-helix
XX dimerisation domain, the human IgG3 hinge domain, the anti-B7.2 MAb 1G10
XX VH region, another flexible linker, the 1G10 VL region and a
XX hexahistidine tag. The B1TAB has been expressed in Escherichia coli
XX cells. It cross-links, and/or cross-reacts, with the costimulatory
XX molecules B7.1 and B7.2 expressed on the membrane of professional antigen
XX -presenting cells, leading to the inhibition of antigen-specific T cell
XX activation. The invention relates to such B7-binding molecules, methods
XX for their production, and their use for treating or preventing diseases
XX of the immune system, in particular graft rejection, graft versus host
XX disease, allergy and autoimmune diseases (claimed)
SQ Sequence 1725 BP; 424 A; 439 C; 468 G; 394 T; 0 U; 0 Other;

Query Match 57.6%; Score 548; DB 2; Length 1725;
Best Local Similarity 86.7%; Pred. No. 2.4e-135;
Matches 631; Conservative 0; Mismatches 85; Indels 12; Gaps 2;

QY 1 ATGGCCGAGGTCAGCTGACAGAGTCAAGGAGGCTTAGTGACCTGGAGGGTCCCGG 60
DB 61 ATGGCCGAGGTCAGCTGACAGAGTCTGGGAGGCTTAGTGACCTGGAGGGTCCCGG 120
QY 61 AAAGTGTCTGAGAGTCCCTGCTGCTTCAGTGGCAGTGGGTTCTGGGACCTTTACTCT 120
DB 121 CCGGCTCTGAGAGTCCCTGCTGCTTCAGTGGCAGTGGGTTCTGGGACCTTTACTCT 180
QY 121 CAGGCTCCAGAGAAAGGGGCTGGAGTGGGTCGCACTATTATTAGTAGGGCAGTAGTACCATC 180
DB 181 CAGGCTCCAGAGAAAGGGGCTGGAGTGGGTCGCACTATTATTAGTAGGGCAGTAGTACCATC 240
QY 181 TACTATGAGACACAGTAGAAGGGACGATTCACATCTCCAGAGACATCCCAAGAACACC 240
DB 241 TACTATGAGACACAGTAGAAGGGACGATTCACATCTCCAGAGACATCCCAAGAACACC 300
QY 241 CTGTTCTGCAAAATAGACAGTCTTAAGTCTTAGAGACAGGTCATGATTTACTGTGCAAGA 300
DB 301 CTGTTCTGCAAAATAGACAGTCTTAAGTCTTAGAGACAGGTCATGATTTACTGTGCAAGA 360
QY 301 GA-----TTACGGGGCTTATTGGGGCCCAAGGACACAGGTCATGATTTACTGTGCA 354
DB 361 GACGGCTGTAATCTTGAATGCTGGGGCCCAAGGACACAGGTCATGATTTACTGTGCA 420
QY 361 GACGGCTGTAATCTTGAATGCTGGGGCCCAAGGACACAGGTCATGATTTACTGTGCA 420
DB 421 GCGGCTCAGGCGAGAGTGGCTCTGGCGGTGGCGGATGAGCATGAGCTCACTCAAGTCT 480

```

|    |  |     |   |     |
|----|--|-----|---|-----|
| Oy |  | 415 | CCAGGAATCAACGTCTGCATATCTCCAGGGAGAAGGTCACCATGACCCTGCAGTGCAGTTCA    | 477 |
| Db |  | 481 | CCAATTCTTAATGCGTGCATCTGTABGGGAAATCGGGTACCATATCACCCTGCAGTGCTCACTCA | 540 |
| Oy |  | 475 | AGTGTAAAGGTACA-----TGAACCTGGTTCACAAGAAATCAGGACACTCCCCCAAAGA       | 528 |
| Db |  | 541 | AGAAATAAATTCCAGCAATTTGGACATGCTATCCAAACAGAAATGCAGAAACTTCCCCTAAAACC | 600 |
| Oy |  | 529 | TGGATTTATGACACATCCAAACTGTCTTTGTGGAAATGCCCTGTCTGCTTCAAGTGGCAATGGG  | 588 |
| Db |  | 601 | TGGATTTATGGAACATCCAACTGGCTTCTGGAAGTCCCTTCTGCTTCAAGTGGCAATGGA      | 660 |
| Oy |  | 589 | TCTGGGACCTCTTACTCTCTCAACAATCAGACAGCATGGAAGTCTGACCTGCAATTAC        | 648 |
| Db |  | 661 | TCTGGGACCGATTATACCTCTCAACAATCAGACAGCATGGAAGTCTGACCTGCAATTAT       | 720 |
| Oy |  | 649 | TACTGCCAGCAGTAGTAGTAATACCACTCACGCTTGGTGGTGGAGCAAGCTGGAGCTG        | 708 |
| Db |  | 721 | TACTGTCAACAGTAGTAGTAATACCACTCACGCTTGGTGGTGGAGCAAGCTGGAGCTG        | 780 |
| Oy |  | 709 | AAACGGGC 716  |     |
| Db |  | 781 | AAACGGAC 788  |     |

RESULT 14  
AAK01652  
ID AAK01652 standard; DNA; 1668 BP.  
XX  
AC AAK01652;  
DT 10-MAY-1999 (first entry)  
XX  
DE Bispesicific tetravalent antibody B17Ab1G10-B7-24H6 DNA.  
XX  
B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2; CD86;  
KM T cell activation; inhibitor; graft versus host disease;  
KM transplant rejection; allograft rejection; autoimmune disease; allergy;  
KM therapy; human; bispesicific tetravalent antibody; B17Ab,  
KW B17Ab1G10-B7-24H6; se.  
XX  
OS Mus sp.  
OS Homo sapiens.  
OS Synthetic.  
OS Chimeric.  
XX  
FN MO9858965-A2.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98MO-EP003791.  
XX  
PR 20-JUN-1997; 97EP-00870092.  
XX  
PA (INNO-) INNOGENETICS NV.  
XX  
PI Lorre K, Sablon E, Buysse M, Bosman A;  
DR WPl; 1999-106615/09.  
DR P-PsDB; AAM90218.  
XX  
PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat  
XX immune diseases including allograft rejection.  
PS Example 7.1; Fig 17, 182pp; English.

This DNA sequence encodes the bispesicific tetravalent antibody B17Ab1G10-B7-24H6 (see AAM90218) comprising the anti-B7.2 murine monoclonal antibody (Mab) 1G10 VH region joined, via a flexible synthetic linker, to the 1G10 VL region, the human IgG3 hinge domain, a helix-turn-helix dimerisation domain, the human IgG3 hinge domain, the anti-B7.1 Mab B7-24 VH region, another flexible linker, the B7-24 VL region and a hexahistidine tag. The

CC B17d has been expressed in *Escherichia coli* cells. It cross-links,  
CC and/or cross-reacts, with the costimulatory molecules B7.1 and B7.2  
CC expressed on the membrane of professional antigen-presenting cells,  
CC leading to the inhibition of antigen-specific T cell activation. The  
CC invention relates to such B7-binding molecules, methods for their  
CC production, and their use for treating or preventing diseases of the  
CC immune system, in particular graft rejection, graft versus host disease,  
CC allergy and autoimmune diseases (claimed)  
XX  
SQ Sequence 1668 BP, 409 A, 424 C, 453 G, 382 T, 0 U, 0 Other;

Query Match Score 539; DB 2; Length 1668;  
Best Local Similarity 86.5%;  
Pred. No. 5.9e-113;  
Matches 622; Conservative 85; Indels 12; Gaps 2

|    |      |   |      |
|----|------|---|------|
| QY | 8    | AGGCTCAAGCTGTCACAGAGTCAAGGGGGAGAGCTTTAGTGCACAGCCGTGAGAGGGCTCCGGAAACCTCT | 67   |
| Db | 932  | AGGTGCAGCTCAAGAGAGTCTGGGGGAGAGCTTAGTGCACAGCCGTGAGAGGGCTCCGGAGGCTCT      | 991  |
| QY | 68   | CCTGTGAGAGCCTCTGGAATTCACTTTCAGTAGCTTTGGAAATGCACTGGGTTGTCAGAGCTC         | 127  |
| Db | 992  | CTGTGTGAGCCTCTGGAATTCACTTTCAGTAGTTTGGAAATGCACTGGGTTGTCAGAGCTC           | 1051 |
| QY | 128  | CAGAGAAAGGAGCTGGAAGTGGGTGCATATATTAAGTAGTGCAGTAGTACCATTAATATG            | 187  |
| Db | 1052 | CAGGAGAGGGAGCTGGATATGGGTGCCATTCAATTAGTAGTGCAGTAGTCTCATTAATATG           | 1111 |
| QY | 188  | CAGACACAGTAGAAGGAGACATTCACCACTTCACAGACAAATCCCAAGAAACCCGTGTCC            | 247  |
| Db | 1112 | CAGACTAGTAGAAGGGCCGATTCACCACTTCACAGACAAAGCCCAAGAAACCCGTGTTC             | 1171 |
| QY | 248  | TGCAATATGACCAAGCTTAAGGCTGAGGACAAGGCATGATTAATCTGTGCAAGAGA-----           | 302  |
| Db | 1172 | TGCAATATGAAACAGTCTAAGGCTGAGGACAAGGCCGATATTAATCTGTGCAAGAGAGCGCT          | 1231 |
| QY | 303  | -TTACGGGAGCTTAATTGAGGAGCCAAAGGACACACGCTACCGTCTCCTCAAGTAGAGGCGAGCT       | 361  |
| Db | 1232 | GGTACTTCGATGTCTGGGGGCCAAAGGACACACGCTCATGTCTCCTCAAGTAGAGGCGGTT           | 1291 |
| QY | 362  | CAGCGGAGAGGTGGCTCTGGCGGTGAGGANTGGAACATTAAGCTCAACCAAGTCTCCAGCA           | 421  |
| Db | 1292 | CAGCGGAGAGGTGGCTCTGGCGGTGAGGANTGGAACATTAAGCTCAACCAAGTCTCCATCTT          | 1351 |
| QY | 422  | TCAATGTCTGCATCTCCAGGGGAGAGGTCACCAATGACTGGAAGTGCACATTAAGTGA              | 481  |
| Db | 1352 | CTAATGTCTGCATCTGTAGGGGAGTGGGTGCACCAATGACTGGAAGTGCACATTAAGTGA            | 1411 |
| QY | 482  | GGTACA-----TGAACGTGGTTCCAAAGAAAGTCAGGACCTCCGCCCAAAAGATGATTT             | 535  |
| Db | 1412 | GTTCCAGCAATTTTGCATCTGTAACCAAGAAAGTCAAGAAATTTCCGCCCAACCCGTGATTT          | 1471 |
| QY | 536  | ATGACACATCCAAACTGTCTTCTGGAAGTCCCTGTGCTTCAATGAGCAGTGGGTCTGGGA            | 595  |
| Db | 1472 | ATGACACATCCAAACTGTGACTTCTGGAAGTCCCTTCTGTGCTTCAATGAGCAGTGGGTCTGGGA       | 1531 |
| QY | 596  | CTCTTAACTCTGCACAAATGACAGCATGAGGCTGAAGATGTGACCTTAATCTAGTCC               | 655  |
| Db | 1532 | CCGATTTATACCTCTCAACATGACAGCATGACAGCAAGATGTGACCTTAATCTAGTCTC             | 1591 |
| QY | 656  | AGCAGTAGAAGTAGAACCACTCACTCACTGGTGTGGGACCAAGCTGGAAGCTGAAGG               | 714  |
| Db | 1592 | AACAGTAGAAGTAGAATATCACTCACTCACTGGTGTAGGGGACCAAGCTCAAGATCAAAAGG          | 1650 |

|         |   |
|---------|---|
| RESULT  | 15  |
| AD58607 |   |
| ID      | AD58607 standard; DNA; 711 BP.              |
| XX      |   |
| AC      | AD58607;                                    |
| DT      |   |
| XX      | 04-DEC-2003 (first entry)                   |
| XX      |   |
| Mouse   | G1 single chain Fv-recombinant antibody DNA |



|                           |   |
|---------------------------|---|
| XX                        | Major histocompatibility complex; MHC; HLA-restricted antigen; cancer;    |
| KW                        | viral infection; autoimmune disease; gene therapy; cytostatic; virucide;  |
| KX                        | immunomodulator; mouse; gene; ds.   |
| OS                        | Mus sp.   |
| XX                        |   |
| FH                        | Key   |
| FT                        | CDS   |
| FT                        | location/Qualifiers   |
| FT                        | 1..711  |
| FT                        | /tag= a   |
| FT                        | /product= "Mouse G1 single chain Fv-recombinant antibody"                 |
| FT                        | /note= "No start and stop codon"  |
| FT                        | /partial  |
| XX                        |   |
| PN                        | WO2003068201-A2.  |
| XX                        |   |
| PD                        | 21-AUG-2003.  |
| XX                        |   |
| PF                        | 11-FEB-2003; 2003WO-IL000105.   |
| XX                        |   |
| PR                        | 13-FEB-2002; 2002US-00073301.   |
| XX                        |   |
| PA                        | (TECR ) TECHNION RES & DEV FOUND LTD.                                     |
| XX                        |   |
| PI                        | Reiter Y, Denkerberg G;   |
| XX                        |   |
| DR                        | WPI; 2003-689603/65.  |
| XX                        |   |
| XX                        | P-PsDB; AAE38657.   |
| PT                        | New isolated molecule comprising an antibody that binds with a human      |
| PT                        | major histocompatibility complex (MHC) class I being complexed with a HLA |
| PT                        | -restricted antigen, useful for treating cancer, viral infection or       |
| XX                        | autoimmune disease.   |
| PS                        |   |
| XX                        | Claim 62; Fig 3a; 81bp; English.  |
| CC                        | The invention relates to an isolated molecule comprising an antibody      |
| CC                        | specifically bindable with a binding affinity below 20 nanomolar to a     |
| CC                        | human major histocompatibility complex (MHC) class I being complexed with |
| CC                        | a HLA-restricted antigen. The molecules, antibodies, and methods are      |
| CC                        | useful for treating cancer, viral infection and an autoimmune disease.    |
| CC                        | The invention is useful in gene therapy. The present sequence is mouse G1 |
| CC                        | single chain Fv-recombinant antibody DNA                                  |
| XX                        |   |
| SQ                        | Sequence 711 BP; 164 A; 176 C; 208 G; 163 T; 0 U; 0 Other;                |
| Query Match               | 56.0%; Score 532.4; DB 9; Length 711;                                     |
| Best Local Similarity     | 85.8%; Pred. No. 2.5e-131;  |
| Matches 609; Conservative | 0; Mismatches 86; Indels 15; Gaps 1                                       |
| OY                        | 8 AGGTCAAGCTGCAGGATCGAGGGGAGGCTTAAAGTCAGCCTTGAGAGGTCGCCGAAAACCTCT 67      |
| Db                        | 2 AGGTGAACCTGCAGGATGCTCGGGGAGGCTTAGTGAAGCCTGGAAGGCTCCGAAAACCTCT 61        |
| OY                        | 68 CCTGTGACGCTCTGTGATTCACTTTCAAGTAAGTTTGAATGCACTGAGTTCTGTCAAGCTC 127      |
| Db                        | 62 CCTGTGACGCTCTGTGATTCACTTTCAAGTAAGTTTGAAGTAAGTTCTGTGAGTTCTGTCAAGCTC 121 |
| OY                        | 128 CAGAGAAGAGGCTGAGTGGTGCAGATAATTAATAGTAGTGCAGTGTAGTACTTACTTANG 187      |
| Db                        | 122 CAGACAAGAGGCTGAGTGGTGCAGAACATTAATAGTAGTGTAGTACTTACTTACTTANG 181       |
| OY                        | 188 CAGACACAGTGAAGGAGCGATTACCATCTCCAGAGACAAATCCCAAAGCACCTGTGTC 247        |
| Db                        | 182 CAGACAGTGTGAAGGGGCGATTACCATCTCCAGAGACAAATCCCAAAGCACCTGTGTC 241        |
| OY                        | 248 TGCAAAATGACCAAGTCTAAGAGTGTGAGAGACGCGTCAATTAATCTGTGTGCAAGAGTACG 307    |
| Db                        | 242 TGCAAAATGACCAAGTCTAAGAGTGTGAGAGACGCGTCAATTAATCTGTGTGCAAGAGTAACT 301   |
| OY                        | 308 GG-----GCTTATTTGGGGGCAAGGGGCAACGGTCAACCGTCTCTCAAGGTC 352              |
| Db                        | 302 GGAAGAGATGTAATCTCAATGCTCTGGGGCAAGGGGCAACGGTCAACCGTCTCTCAAGGTC 361     |

Oy 353 AAGGCGGCTCAAGCGGAGGTGGCTTGGCGGTGGGAGGATTCGAACTTGAAGCTCAACCAGT 412  
 Db 362 GAGGCGGTTCAGGCGGAGGTGGCTTGGCGGTGGGAGGATTCGAACTCGAGCTCACTCAGT 421  
 Oy 413 CTCGAGCAATCATGTCTGCACTCTCCAGGGGAGAGGGTCAACCATGACTCTGCAGTCCAGTT 472  
 Db 422 CTCGAGCAATCATGTCTGCACTCTCCAGGGGAGAGGGTCAACCATGACTCTGCAGTCCAGTT 481  
 Oy 473 CAAGTGAAGTACATGAATCGTGTCCAAAGAGTCAAGGCACTCCCAAAAGATGGA 532  
 Db 482 CAAGTATACGTACATATATTGGTACCAAGAGGCTTGATCTCTCCCAAGCTCCTGA 541  
 Oy 533 TTATAGACAATCCAACTGTCTTTGGAAGTCCCTGCTCGCTTCAATGGCAGTGGGTCTG 592  
 Db 542 TTATAGACAATCCAACTGTGCTCTGGAAGTCCCTTTTCGCTTCAATGGCAGTGGGTCTG 601  
 Oy 593 GGACCTCTTAATCTCTCAAACTACAGACAGATGAGGCTGAAAGTGTGCACTTAATCT 652  
 Db 602 GGACCTCTTAATCTCTCAAACTACAGACAGATGAGGCTGAAAGTGTGCACTTAATCT 661  
 Oy 653 GCCAGCAATGAGTAGTAACCACTCAACAGTTCGGGTCTGGGACCAAGCTG 702  
 Db 662 GCCAGCAATGAGTAGTAACCACTCAACAGTTCGGGACCAAGCTG 711

Search completed: June 23, 2005, 07:10:00  
Job time : 621 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: June 23, 2005, 09:03:52 ; Search time 689 Seconds  
(without alignments)  
8627.347 Million cell updates/sec

Title: US-09-403-882a-1  
Perfect score: 951  
Sequence: 1 atggccgaggtcagctgca.....ggcagaagaagcagctag 951

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues  
Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
12: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
17: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
18: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
19: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
20: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
21: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq.\*  
22: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq.\*  
24: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 949   | 99.8        | 8902   | 17    | US-10-221-276-5   |
| 2          | 460.2 | 48.0        | 807    | 16    | US-10-408-930-4   |
| 3          | 459.6 | 48.3        | 873    | 18    | US-10-406-830-36  |
| 4          | 448.8 | 47.2        | 726    | 15    | US-10-259-087A-19 |
| 5          | 448.8 | 47.2        | 726    | 19    | US-10-689-006-19  |
| 6          | 446   | 46.9        | 861    | 18    | US-10-406-830-31  |
| 7          | 443.6 | 46.6        | 1176   | 15    | US-10-075-947A-3  |

| 8  | 442   | 46.5  | 717  | 9  | US-09-808-037-5   | Sequence 5, Appl1   |
|----|-------|---|------|----|-------------------|---------------------|
| 9  | 442   | 46.5 <td>717</td> <td>14</td> <td>US-10-162-889-5</td> <td>Sequence 5, Appl1</td> | 717  | 14 | US-10-162-889-5   | Sequence 5, Appl1   |
| 10 | 442   | 46.5 <td>717</td> <td>17</td> <td>US-10-384-788-5</td> <td>Sequence 5, Appl1</td> | 717  | 17 | US-10-384-788-5   | Sequence 5, Appl1   |
| 11 | 442   | 46.5 <td>717</td> <td>18</td> <td>US-10-618-856-5</td> <td>Sequence 5, Appl1</td> | 717  | 18 | US-10-618-856-5   | Sequence 5, Appl1   |
| 12 | 442   | 46.5 <td>717</td> <td>21</td> <td>US-10-749-522-5</td> <td>Sequence 5, Appl1</td> | 717  | 21 | US-10-749-522-5   | Sequence 5, Appl1   |
| 13 | 438.2 | 46.1  | 720  | 9  | US-09-976-787-30  | Sequence 30, Appl1  |
| 14 | 436   | 45.8  | 780  | 4  | US-10-247-488-3   | Sequence 3, Appl1   |
| 15 | 434.8 | 45.7  | 717  | 18 | US-08-940-544-3   | Sequence 3, Appl1   |
| 16 | 434.8 | 45.7  | 717  | 15 | US-10-075-947A-1  | Sequence 1, Appl1   |
| 17 | 434.2 | 45.7  | 717  | 21 | US-10-688-255-6   | Sequence 6, Appl1   |
| 18 | 434   | 45.6  | 873  | 18 | US-10-406-830-35  | Sequence 35, Appl1  |
| 19 | 433.8 | 45.6  | 714  | 10 | US-09-798-689-22  | Sequence 22, Appl1  |
| 20 | 430.6 | 45.3  | 714  | 9  | US-09-976-787-31  | Sequence 31, Appl1  |
| 21 | 430.6 | 45.3  | 711  | 14 | US-10-247-488-1   | Sequence 1, Appl1   |
| 22 | 428.6 | 45.1  | 786  | 19 | US-10-689-006-23  | Sequence 23, Appl1  |
| 23 | 427.6 | 45.0  | 746  | 17 | US-10-354-246-3   | Sequence 3, Appl1   |
| 24 | 426.8 | 44.9  | 1673 | 10 | US-09-819-266-27  | Sequence 27, Appl1  |
| 25 | 424.2 | 44.6  | 723  | 18 | US-10-062-188-4   | Sequence 4, Appl1   |
| 26 | 422   | 44.4  | 732  | 21 | US-10-879-994-15  | Sequence 15, Appl1  |
| 27 | 422   | 44.4  | 732  | 21 | US-10-610-452-15  | Sequence 15, Appl1  |
| 28 | 421.6 | 44.3  | 726  | 15 | US-10-259-087A-17 | Sequence 17, Appl1  |
| 29 | 421.6 | 44.3  | 726  | 19 | US-10-689-006-17  | Sequence 17, Appl1  |
| 30 | 421.2 | 44.3  | 738  | 21 | US-10-861-617-14  | Sequence 14, Appl1  |
| 31 | 417.4 | 43.9  | 768  | 18 | US-10-239-656-60  | Sequence 60, Appl1  |
| 32 | 417.4 | 43.9  | 1509 | 18 | US-10-239-656-76  | Sequence 76, Appl1  |
| 33 | 416.2 | 43.8  | 729  | 17 | US-10-354-246-6   | Sequence 6, Appl1   |
| 34 | 410   | 43.1  | 720  | 21 | US-10-935-290-195 | Sequence 195, Appl1 |
| 35 | 410   | 43.0  | 738  | 21 | US-10-861-617-16  | Sequence 16, Appl1  |
| 36 | 409.2 | 43.0  | 717  | 21 | US-10-935-290-145 | Sequence 145, Appl1 |
| 37 | 407   | 42.8  | 756  | 16 | US-10-169-351-68  | Sequence 68, Appl1  |
| 38 | 406.6 | 42.8  | 924  | 9  | US-09-742-693-26  | Sequence 26, Appl1  |
| 39 | 405.6 | 42.6  | 728  | 20 | US-10-798-845-30  | Sequence 30, Appl1  |
| 40 | 404.8 | 42.6  | 1479 | 19 | US-10-682-845-66  | Sequence 66, Appl1  |
| 41 | 404.8 | 42.6  | 1479 | 19 | US-10-682-845-76  | Sequence 76, Appl1  |
| 42 | 403.2 | 42.4  | 1241 | 18 | US-10-362-591-1   | Sequence 1, Appl1   |
| 43 | 403.2 | 42.4  | 1479 | 19 | US-10-682-845-60  | Sequence 60, Appl1  |
| 44 | 403.2 | 42.4  | 1479 | 19 | US-10-682-845-62  | Sequence 62, Appl1  |
| 45 | 403.2 | 42.4  | 1479 | 19 | US-10-682-845-68  | Sequence 68, Appl1  |

## ALIGNMENTS

RESULT 1  
US-10-221-276-5  
; Sequence 5, Application US/10221276  
; Publication No. US20030211495A1  
; GENERAL INFORMATION:  
; APPLICANT: Tivv Telethon Institute for Child Health Research  
; TITLE OF INVENTION: Improved Reverse Two Hybrid Screening Method  
; FILE REFERENCE: multiple reporter pct.doc  
; CURRENT APPLICATION NUMBER: US/10/221,276  
; CURRENT FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 8902  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: VECTOR  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1).....(8902)  
; OTHER INFORMATION: n = A, T, C or G  
; US-10-221-276-5

Query Match 99.8%; Score 949; DB 17; Length 8902;  
Best Local Similarity 99.8%; Pred. No. 1.7e-281;  
Matches 949; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 ATGGCCGAGGTCAAGCTGACAGAGTCAAGGAGGCTTATGTCAGGCTGAGGCTCCCG 60  
|||||

```

Db      2820 ATGGCCGAGTCAAGTCGACGAGTACAGGGGAGGCTTGTGTCAGCCTTGAGGGCTCCCG 2879
Qy      61 AAACCTCTCTGTGACACCTCTGTGATTCACCTTCAGTAGCTTTGGAATGCACTGGTTCGT 120
Db      2880 AAACCTCTCTGTGACACCTCTGTGATTCACCTTCAGTAGCTTTGGAATGCACTGGTTCGT 2939
Qy      121 CAGGCTCCAGAGAGAGGGGCTGAGTGGGTCCGATATATTAGTAGGCGTAGTACATC 180
Db      2940 CAGGCTCCAGAGAGAGGGGCTGAGTGGGTCCGATATATTAGTAGGCGTAGTACATC 2999
Qy      181 TACTATGACAGACACAGTGAAGGGACGATTCACCATCTCCAGAGACAATCCCAAGAACAC 240
Db      3000 TACTATGACAGACACAGTGAAGGGACGATTCACCATCTCCAGAGACAATCCCAAGAACAC 3059
Qy      241 CTGTTCCTGCAATGACAGTCTAAGTCTGAGGACACGATCATGTTATTAATCTGTGCAGA 300
Db      3060 CTGTTCCTGCAATGACAGTCTAAGTCTGAGGACACGATCATGTTATTAATCTGTGCAGA 3119
Qy      301 GATTACGGGGCTTATTGGGGCCAGAGGACACGATCACTGTCTCTCAGGTGAGGCGGC 360
Db      3120 GATTACGGGGCTTATTGGGGCCAGAGGACACGATCACTGTCTCTCAGGTGAGGCGGC 3179
Qy      361 TCAGGCGAGGTGGCTCTGGCGGTGGGATGGGACATTTGAGCTCACCGAGTCTCCAGCA 420
Db      3180 TCAGGCGAGGTGGCTCTGGCGGTGGGATGGGACATTTGAGCTCACCGAGTCTCCAGCA 3239
Qy      421 ATCATGTCTGCACTCTCAGAGGGAGAGGATCACCATGACCTGCAGTGCAGTTCAAGTGA 480
Db      3240 ATCATGTCTGCACTCTCAGAGGGAGAGGATCACCATGACCTGCAGTGCAGTTCAAGTGA 3299
Qy      481 AGGTACATGAACTGTGTTCCACAGAAATCAGAGCACTCCCCCAAAAGATGATTTATGAC 540
Db      3300 AGGTACATGAACTGTGTTCCACAGAAATCAGAGCACTCCCCCAAAAGATGATTTATGAC 3359
Qy      541 ACATCCAACTGTCTCTGGAAGTCCCTGCTGCTGAGTGGGAGTGGGATCTGAGGACCTCT 600
Db      3360 ACATCCAACTGTCTCTGGAAGTCCCTGCTGCTGAGTGGGAGTGGGATCTGAGGACCTCT 3419
Qy      601 TACTCTCTCACAATCAGACATGAGAGGCTGAAGATGTGCCACTTACTACTGCCAGAC 660
Db      3420 TACTCTCTCACAATCAGACATGAGAGGCTGAAGATGTGCCACTTACTACTGCCAGAC 3479
Qy      661 TGGAGTAGTAACTCCACTCAGTTCGGTGTCTGGGACCAAGCTGGAAGTGAAGCGGCGGCC 720
Db      3480 TGGAGTAGTAACTCCACTCAGTTCGGTGTCTGGGACCAAGCTGGAAGTGAAGCGGCGGCC 3539
Qy      721 GCAGAACAAAATCTATCTCAGAGAGGATCTGAATGGGGCCGTGCAGCAACAAAATCTC 780
Db      3540 GCAGAACAAAATCTATCTCAGAGAGGATCTGAATGGGGCCGTGCAGCAACAAAATCTC 3599
Qy      781 ATCTCAGAAAGAGATCTGAATGCTGTGGCCAGGACACGACAGAGGATCATGCTGTGCCA 840
Db      3600 ATCTCAGAAAGAGATCTGAATGCTGTGGCCAGGACACGACAGAGGATCATGCTGTGCCA 3659
Qy      841 CACTCTCTGCGCTTTAAGTGGTGGTGAATCTCAGCCATCTTGCGCTGTGGTCTCAC 900
Db      3660 CACTCTCTGCGCTTTAAGTGGTGGTGAATCTCAGCCATCTTGCGCTGTGGTCTCAC 3719
Qy      901 ATCATCTCCCTTATCATCTCTCATCATGCTTTGGCAGAAAGCAAGCTTAA 951
Db      3720 ATCATCTCCCTTATCATCTCTCATCATGCTTTGGCAGAAAGCAAGCTTAA 3770

```

```

; PRIOR APPLICATION NUMBER: PCT/AU99/00341
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 09/674,677
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: AU P23445
; PRIOR FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(807)
; OTHER INFORMATION: Sequence of the anti-hepatitis surface antigen (4c2) scFv
us-10-408-930-4

Query Match      48.4%; Score 460.2; DB 16; Length 807;
Best Local Similarity 77.5%; Pred. No. 4e-131;
Matches 574; Conservative 0; Mismatches 158; Indels 9; Gaps 1;

1 ATGGCCGAGTCAAGTCGACGAGTACAGGGGAGGCTTGTGTCAGCCTTGAGGGCTCCCG 60
3 ATGGCCGAGTCAAGTCGACGAGTACAGGGGAGGCTTGTGTCAGCCTTGAGGGCTCCCG 62
61 AAACCTCTCTGTGACACCTCTGTGATTCACCTTCAGTAGCTTTGGAATGCACTGGTTCGT 120
63 AAAGTACTCTGCAATGACAGTCTAAGTCTGAGGACACGATCATGTTATTAATCTGTGCAGA 122
121 CAGGCTCCAGAGAGAGGGGCTGAGTGGGTCCGATATATTAGTAGGCGTAGTACATC 180
123 CAGAGTCAATGCGCAAGTCTAAGTGAATGACATTAATTAATCTTTGGTATATCA 182
181 TACTATGACAGACACAGTGAAGGGACGATTCACCATCTCCAGAGACAATCCCAAGAACAC 240
183 AACTCAACCAAGATTTGAGGACCAAGCCACATATGATCTGACAAATCTCCACACACA 242
241 CTGTTCCTGCAATGACAGTCTAAGTCTGAGGACACGATCATGTTATTAATCTGTGCAGA 300
243 GGCTATTTGGAACCTTGCGAGATGACATCTGAGGATTTGCGCATCTTACTGTGCAGAA 302
301 G-----ATTACGGGGCTTATTGGGGCCAGAGGACACGATCACTGTCTCTCAGGT 351
303 GTGATGCACTGGTCTTTCGATGTCTGGGGCCAGAGGACACGATCACTGTCTCTCAGGT 362
352 GGAGGCGGCTCAGAGGAGGTGGCTCTGGCGGTGGGATGGGACATTTGAGGCTCAACCCAG 411
363 GGAGGCGGCTCAGAGGAGGTGGCTCTGGCGGTGGGATGGGACATTTGAGGCTCAACCCAG 422
412 TCTCAGCAATCATGTCTGCAATCTCCAGGGAGAGGATCACCATGACCTGCAGTCCAGT 471
423 TCTCAGCAATCATGTCTGCAATCTCCAGGGAGAGGATCACCATGACCTGCAGTCCAGT 482
472 TCAAGTGAAGTACATGAATCTGTTCCAGAGAGTCAAGGCACTCCCCCAAAAGATGG 531
483 TCACCTGTCAAGTACATGACATCTGATCCACAGAGTCAAGGCACTCCCCCAAAAGATGG 542
532 ATTATGACACATCCAACTGTCTTGGAGTCCCTGCTGCTGCTTGAAGTGGAGTGGGTCT 591
543 ATTATGACACATCCAACTGTCTTGGAGTCCCTGCTGCTGCTTGAAGTGGAGTGGGTCT 602
592 GGAGCTCTTACTCTCTCAGATCAGCAGATGAGGCTGAAGATGCTGCCACTTATAC 651
603 GGAGCTCTTACTCTCTCAGATCAGCAGATGAGGCTGAAGATGCTGCCACTTATAC 662
652 TGCCAGCAGTGAAGTGAATCCACTCAGTTCGGTGTCTGGGACCAAGTGAAGTGAAGTGA 711
663 TGCCAGCAGTGAAGTGAATCCACTCAGTTCGGTGTCTGGGACCAAGTGAAGTGAAGTGA 722
712 CGGGGCGGCGCAGAACAAA 732
723 CGGGGCGGCGCAGATTATAAA 743

```

```

; Sequence 4, Application US/10408930
; Publication No. US20030170820A1
; GENERAL INFORMATION:
; APPLICANT: Coia, et al.
; TITLE OF INVENTION: CONTINUOUS IN-VITRO EVOLUTION
; FILE REFERENCE: 674537-2003.1
; CURRENT APPLICATION NUMBER: US/10/408,930
; CURRENT FILING DATE: 2003-04-07

```

RESULT 3  
US-10-406-830-36  
Sequence 36, Application US/10406830  
Publication No. US20040071696A1  
GENERAL INFORMATION:  
APPLICANT: ADAMS, GREGORY P.  
APPLICANT: HORAK, EVA M.  
APPLICANT: WEINER, LOUIS M.  
APPLICANT: JAMES, MARKS D.  
TITLE OF INVENTION: BISPECTIFIC SINGLE CHAIN Fv ANTIBODY MOLECULES AND METHODS OF USE  
FILE REFERENCE: 407T-000410US  
CURRENT APPLICATION NUMBER: US/10/406,830  
CURRENT FILING DATE: 2003-04-04  
PRIOR APPLICATION NUMBER: US 60/370,276  
PRIOR FILING DATE: 2002-04-05  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 36  
LENGTH: 873  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Nucleic acid encoding single chain antibody.  
US-10-406-830-36

Query Match 48.3%; Score 459.6; DB 18; Length 873;  
Best Local Similarity 75.4%; Pred. No. 6,4e-131;  
Matches 611; Conservative 0; Mismatches 169; Indels 30; Gaps 2;

QY 1 ATGGCCGAGGTCAAGTGCAGAGTCAAGGGGGGCTTATGAGCTGGAGGGTCCCG 60  
DB 61 ATGGCCGAGGTCAAGTGCAGAGTGCAGGGGGGAGGATGTCCAGCTGGAGGGTCCCG 120  
QY 61 AAATCTCTCTGTCAGAGCTCTGATTCATTCTAGTACTTGAATGCACTGGTTCGT 120  
DB 121 AGACTCTCTCTGTCAGAGCTCTGATTCATTCTAGTACTTGAATGCACTGGTTCGT 180  
QY 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCAATATTATGATGAGTGGCACTAGTACC 180  
DB 181 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCAATATTATGATGAGTGGTGTAGCACA 240  
QY 181 TACTATGACAGACACTGTAAGGAGGACATTCACATCTCCAGAGACATCCCAAGAAC 240  
DB 241 TACTATGACAGACACTGTAAGGAGGACATTCACATCTCCAGAGACATCCCAAGAAC 300  
QY 241 CTGTTCTCTGCAATGACAGTCTAAGTCTGAGGACACAGGTCACTGATTAATGTCAGA 300  
DB 301 CTGTTCTCTGCAATGACAGTCTAAGTCTGAGGACACAGGTCACTGATTAATGTCAGA 360  
QY 301 GA-----TTACGGGGCTTATTGGGGCCAAAGGACCAAG 333  
DB 361 GAGGAGTATGACAGCAATGGAATTAATGTAATCTGATCTTGGGGCCGTGGCAACCTG 420  
QY 334 GTTACCTCTCTCTCAAGTGCAGAGGCGCTCAGGGGAGGTGGCTCTGGCGGTGGCGATCG 333  
DB 421 GTTACCTCTCTCTCAAGTGCAGAGGCGCTCAGGGGAGGTGGCTCTGGCGGTGGCGATCG 480  
QY 394 GACATTGAGTCAACCGAGTCTCCAGCAATCATGTCTGATCTCCAGGGGAGAGGGTACC 453  
DB 481 GAAATTGTGTGACTCATGCTCTCTCTCTGATCTCTGAGGAGACAGATGAC 540  
QY 454 ATGACCTGAGTGCAGTTCAG--TGTAAGTACATGAATGTTTCAACAGAACTCA 510  
DB 541 ATCATCTGCGGGCAAGTCAAGACATTAAGCATTAATTAATGTAATGTAATGTAATGTA 600  
QY 511 GGGACCTCCCCCAAAAGATGATTAATGACATCCAACTGCTTCTGAGTCCCTGCT 570  
DB 601 GGGAAAGCCCTTAAGCTCTGATCTTAATGTCATCCAGTTTCAAAAGTGGGGTCCCATCA 660  
QY 571 CGCTTCAAGTGGAGTGGGTCTGGGACCTTACTCTCTCAAAATGACAGCATGAGGCT 630

DB 661 AGGTTCAAGTGGAGTGGATCTGGGACAGATTTCACTCTCACTCAGCATGAGCCCTGCGGCT 720  
QY 631 GAAGATCTGCCACTTACTACTGCGCAGAGTGGAGTGAATTAACCACTCACTGCTGGGCT 690  
DB 721 GACGATTTGCAACTTATTAATCTGCAAGATTAATTAATTAATTAATTAATTAATTAATTA 780  
QY 691 GGGACCAAGCTGGAGTGAAGAGGGGCGCGCAGAACAAACTCATCTCAGAAAGAGAT 750  
DB 781 GGGACCAAGCTGGAGTGAAGATCAAGTGGGCGCGCAGAACAAACTCATCTCAGAAAGAGAT 840  
QY 751 CTGAATGGGGCGCTGACACAAACAACTC 780  
DB 841 CTGAATGGGGCGCGCATCATCATCATCATC 870

RESULT 4  
US-10-259-087A-19  
Sequence 19, Application US/10259087A  
Publication No. US20030130190A1  
GENERAL INFORMATION:  
APPLICANT: Vanderbilt University  
APPLICANT: Hallahan, Dennis E  
APPLICANT: Qu, Shimlan  
TITLE OF INVENTION: IN VIVO PANNING FOR LIGANDS TO RADIATION-INDUCED MOLECULES  
FILE REFERENCE: 1242/47/2  
CURRENT APPLICATION NUMBER: US/10/259,087A  
CURRENT FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: US 60/328123  
PRIOR FILING DATE: 2001-10-03  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19  
LENGTH: 726  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificial antibody ligand number 2  
NAME/KEY: CDS  
LOCATION: (1)..(726)  
OTHER INFORMATION:  
US-10-259-087A-19

Query Match 47.2%; Score 448.8; DB 15; Length 726;  
Best Local Similarity 77.4%; Pred. No. 1.3e-127;  
Matches 562; Conservative 0; Mismatches 152; Indels 12; Gaps 1;

QY 1 ATGGCCGAGGTCAAGTGCAGAGTCAAGGGGAGGCTTATGAGCTGGAGGGTCCCG 60  
DB 1 ATGGCCGAGGTCAAGTGCAGAGTCAAGGGGAGGCTTATGAGCTGGAGGGTCCCG 60  
QY 61 AAATCTCTCTGTCAGAGCTCTGATTCATTCTAGTACTTGAATGCACTGGTTCGT 120  
DB 61 AAATCTCTCTGTCAGAGCTCTGATTCATTCTAGTACTTGAATGCACTGGTTCGT 120  
QY 61 AAGATGCTCTGCAAGGCTCTGATTAATCACTCACTAGTATGATGTAATGTAATGTAATGTA 120  
DB 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCAATATTATGATGAGTGGCACTAGTACC 180  
QY 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCAATATTATGATGAGTGGCACTAGTACC 180  
DB 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCAATATTATGATGAGTGGCACTAGTACC 180  
QY 181 TACTATGACAGACACTGTAAGGAGGACATTCACATCTCCAGAGACATCCCAAGAAC 240  
DB 181 AAGTATCAATGAGAGTCAAGGACAGGCGCACTGATTCAGACAAATCTCCAGACACA 240  
QY 241 CTGTTCTCTGCAATGACAGTCTAAGTCTGAGGACACGCTCATGTAATTAATGTCAGA 300  
DB 241 GCTTATGAGAGTCTCAGAGCTGACCTCTGAGAGCTCTGCGGTCTAATTAATGTCAGA 300  
QY 301 -----GATTAAGGGGCTTATTGGGGCCAAAGGACCAAGGTCAACGCTCTCA 348  
DB 301 TTTGTAATCACTAGCTCTTGAATCTTATGAGGCGCAAGGAGCCAGCGGTCTCTCA 360  
QY 349 GGTGAGAGCGGCTCAGGCGAGGTGGCTCTGGCGGTGGGATGGACATTTGAGTCAACC 408

Db 361 GGTGAGGCGGTTTCAGGCGGAGGTGCTGTGGCGGTGGGATCGGACATTGAGCTCACCC 420  
Qy 409 CAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAGAGGGGTCAACATGACCTGCAGTGGCC 468  
Db 421 CAGTCTCCAAACATCATGTCTGCATCTCCAGGGGAGAGAGTCAACATGACCTGCAGTGGCC 480  
Qy 469 AGTTCAAGTGAAGTGAATGTAAGTGTTCGCAACAGAGTCAAGGACCTCCGCCAAAGAA 528  
Db 481 AGCTCAAGTGAAGTGAATGTAAGTGTTCGCAACAGAGTCAAGGACCTCCGCCAAAGAA 540  
Qy 529 TGGATTATGACACATCCAAAGTCTTGTGAGTCCCTGCTCGCTTCAAGTGGCACTGGG 588  
Db 541 TGGATTATGACACATCCAAAGTCTTGTGAGTCCCTGCTCGCTTCAAGTGGCACTGGG 600  
Qy 589 TCTGGGACCTCTTACTCTCTCAACATCAGACAGATGAGAGCTGAAGTCCCACTTAC 648  
Db 601 TCTGGGACCTCTTACTCTCTCAACATCAGACAGATGAGAGCTGAAGTCCCACTTAT 660  
Qy 649 TACTGCCAGAGTGAAGTGAATGTAACCACTCAAGTTCGTGGGACCAAGCTGGAGCTG 708  
Db 661 TACTGTCAACAGTGAAGTGAATGTAACCACTCAAGTTCGTGGGAGGAGGAGCAAGCTGAATA 720  
Qy 709 AAACGG 714  
Db 721 AAACGG 726

RESULT 5  
US-10-689-006-19  
; Sequence 19, Application US/10689006  
; Publication No. US20040191249A1  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Hallahan, Dennis E  
; APPLICANT: Mernaugh, Raymond  
; TITLE OF INVENTION: PHAGE ANTIBODIES TO RADIATION-INDUCIBLE NEOANTIGENS  
; FILE REFERENCE: 1242/72  
; CURRENT APPLICATION NUMBER: US/10/689,006  
; CURRENT FILING DATE: 2003-10-20  
; PRIOR APPLICATION NUMBER: US 09/914,605  
; PRIOR FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: US 10/259,087  
; PRIOR FILING DATE: 2002-09-27  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 19  
; LENGTH: 726  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Artificial antibody ligand number 2  
; NAME/KEY: CDS  
; LOCATION: (1)..(726)  
US-10-689-006-19

Query Match 47.2%; Score 448.8; DB 19; Length 726;  
Best Local Similarity 77.4%; Pred. No. 1.3e-127;  
Matches 562; Conservative 0; Mismatches 152; Indels 12; Gaps 1;

Qy 1 ATGGCCGAGGTCAAGCTGACAGAGTCAGGGGAGGCTTAGTGCAGCCTGAGGGGTCCGG 60  
Db 1 ATGGCCGAGGTCAAGCTGACAGAGTCAGGGGAGGCTTAGTGCAGCCTGAGGGGTCCAGTG 60  
Qy 61 AAACCTCTCTGTGCAAGCCTCTGATTCATCTTCAAGAGCTTGAATGACAGCTGGTTCCT 120  
Db 61 AAAGATGTCTGTGAAGGCTTCTGATACATTCACATTCAGTATGATGACAGCTGGGTGAAG 120  
Qy 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCATATATAGTAGTGGCAGTAGTACCATC 180  
Db 121 CAGGAAGCTTGGGCGAGGCTTGAAGTGAATTGATATATTAATCTTACATATAGTACT 180

Qy 181 TACTATGACACACAGTGAAGGAGCATTCACATCTCCAGAGACATCCCAAGAACCC 240  
Db 181 AAGTCAATGAGAAAGTTCAAAAGGAGGCCACATGATTCGACAAATCTCCAGACACA 240  
Qy 241 CTGTTCTGCAAAATGACAGTCTAAGTCTGAGGACACGATCATGTATTAATCTGTCAGAA 300  
Db 241 GCCTACATGAGAGCTCAGAGGCTGACCTCTGAGAGCTTGGGGTATTATTAATCTGTCAGAA 300  
Qy 301 -----GATTACGGGCGCTTATTTGGGCGCAAGGAGCCACCGTCAACCGTCTCTCA 348  
Db 301 TTTGGTAACTACGGGCTTTTGGGACCTACGAGGAGCCAAAGGAGCCACCGTCAACCGTCTCTCA 360  
Qy 349 GGTGAGGCGGCTCAGAGGAGAGTGGCTCTGCGGTGGCGGATGGACATTGAGCTCACCC 408  
Db 361 GGTGAGGCGGTTCAAGGGGAGAGTGGCTCTGCGGTGGCGGATGGACATTGAGCTCACCC 420  
Qy 409 CAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAGAGGTCACATGACCTGCAGTGGCC 468  
Db 421 CAGTCTCCAAACATCATGTCTGCATCTCCAGGGGAGAGGTCACATGACCTGCAGTGGCC 480  
Qy 469 AGTTCAAGTGAAGTGAATGTAAGTGTTCGCAACAGAGTCAAGGACCTCCGCCAAAGAA 528  
Db 481 AGCTCAAGTGAAGTGAATGTAAGTGTTCGCAACAGAGTCAAGGACCTCCGCCAAAGAA 540  
Qy 529 TGGATTATGACACATCCAAAGTCTTGTGAGTCCCTGCTCGCTTCAAGTGGCACTGGG 588  
Db 541 TGGATTATGACACATCCAAAGTCTTGTGAGTCCCTGCTCGCTTCAAGTGGCACTGGG 600  
Qy 589 TCTGGGACCTCTTACTCTCTCAACATCAGACAGATGAGAGCTGAAGTCCCACTTAC 648  
Db 601 TCTGGGACCTCTTACTCTCTCAACATCAGACAGATGAGAGCTGAAGTCCCACTTAT 660  
Qy 649 TACTGCCAGAGTGAAGTGAATGTAACCACTCAAGTTCGTGGGACCAAGCTGGAGCTG 708  
Db 661 TACTGTCAACAGTGAAGTGAATGTAACCACTCAAGTTCGTGGGAGGAGGAGCAAGCTGAATA 720  
Qy 709 AAACGG 714  
Db 721 AAACGG 726

RESULT 6  
US-10-406-830-31  
; Sequence 31, Application US/10406830  
; Publication No. US20040071696A1  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, GREGORY P.  
; APPLICANT: HORAK, EVA M.  
; APPLICANT: WEINER, LOUIS M.  
; TITLE OF INVENTION: BISPECIFIC SINGLE CHAIN Fv ANTIBODY MOLECULES AND METHODS OF USE  
; FILE REFERENCE: 407T-000410US  
; CURRENT APPLICATION NUMBER: US/10/406,830  
; CURRENT FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: US 60/370,276  
; PRIOR FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 31  
; LENGTH: 861  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Nucleic acid encoding single chain antibody.  
US-10-406-830-31

Query Match 46.9%; Score 446; DB 18; Length 861;  
Best Local Similarity 74.6%; Pred. No. 9.9e-127;  
Matches 595; Conservative 0; Mismatches 185; Indels 18; Gaps 2;

Qy 1 ATGGCCGAGGTCAAGCTGACAGAGTCAGGGGAGGCTTAGTGCAGCCTGAGGGGTCCGG 60

```

Db      61  ATGGCCAGATGCACAGTGGATGAGTCTGGGAGGAGGGGTGGTCCAGGCTGGGAAGTCCCTG 120
Qy      61  AAACCTCTCTGTGCAGGCTCTGTGATTCACCTTTACAGTACTTTGGAAATGCACTGGGTTGGT 120
Db      121  AGACTCTCTGTGCAGGCTCTGTGATTCACCTTACAGTACTATATATATACCTGGTCCGC 180
Qy      121  CAGGCTCCAGAAAGGGGCTGGAGTGGGTGGCATATATATAGTATGGAGTACTACCATC 180
Db      181  CAGGCTCCAGGCAAGGGGCTGGAGTGGATGGCAGTTATTTCAATATATGGCAATTAATAA 240
Qy      181  TACTATGACAGACACAGTGAAGGGACGATTCACCATCTCCAGAGACAAATCCCAAGAACCC 240
Db      241  TACTACGCGCGCTCCGTGTAAGGACCGATTCACCATCTCCAGAGACAAATCCCAAGAACG 300
Qy      241  CTGTTCTCTGCAAAATGACACAGTCTTAAGGTCTGAGGACACGGTCATGTAATTACTGTGA 300
Db      301  GTGTCTCTGCAAAATGAACAGCTGAGAGCTGAGGACACGGCTGTGTAATTACTGTGAGAG 360
Qy      301  GATTACGGGGCT-----TATTGGGGGCAAGGACCAACGGTCAACCGTCTCC 345
Db      361  GATCTCTACGGGACTACGCTCTTTGACTATCTGGGGGCAAGGAAACCTGTGACACCGTCTCC 420
Qy      346  TCAGGTGAGGCGGCTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATTCGACATTTGAGCTC 405
Db      421  TCAGGTGAGGCGGCTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATTCGACATTCAGATG 480
Qy      406  ACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGAGAGGGTCAACATGACCTGCAGT 465
Db      481  ACCCAGTCTCTTCCACACCTGTCTGCATCTCTGGGAGACAGAGTCAACATCATCTTGCCGG 540
Qy      466  GCCAGTTCAAG---TGTAAGGTACATGAACGTGTTCCACAGAAATCAGGACCTGCCCC 522
Db      541  GCCATCTCAGAGTATGTGTAGCTGTGGTGGCTGGTATCACACAGAAACAGGGAAAGCCCT 600
Qy      523  AAAAGATGGATTTATGACACATCCAAACTGTCTTTCGAGTCCCTGCTGCCTCAGTGGC 582
Db      601  AAACCTCTGATCTATATAGGCGCTCTACTTTAGAAATGGGGTCCCATCAAGGTTTACCGGC 660
Qy      583  AGTGGTCTGGGACCTCTTACTCTCTCACAATTCAGCAGCATGGAGGCTGAAATGTGTGCC 642
Db      661  AGTGATCTGGGACAGAAATTACTCTCACAATTCAGCGGCTCCAGCTGAAAGATTTTGGCA 720
Qy      643  ACTTACTACTGCGCAGCAGTGGAGTGTAAACCACTACGTTGGTCTCTGGACCAAGCTG 702
Db      721  ACTTATTACTGTCAAGAGCTTGTGTAGTTAACCGCTCACTTTCGGCCGAGGGACCAAGGTG 780
Qy      703  GAGCGTGAAGCGGGCGGCGCAGAAACAAAACTCATCTCAGAAAGAGATCTGAATGGGGCC 762
Db      781  GAATTCAAACGTGGGCGCGCAGAAACAAAACTCATCTCAGAAAGAGATCTGAATGGGGCC 840
Qy      763  GTCGACGAACAAAATCTC 780
Db      841  GCACATCAACATCATCAC 858

RESULT 7
US-10-075-947A-3
: Sequence 3, Application US//10075947A
: Publication No. US20030147808A1
: GENERAL INFORMATION:
: APPLICANT: Cheung, Nai-Kong V.
: APPLICANT: Larson, Steven M.
: APPLICANT: Guo, Hong-Fen
: APPLICANT: Rivlin, Ken
: APPLICANT: Sadelain, Michel
: TITLE OF INVENTION: Single Chain FV Constructs of Anti-Ganglioside GD2
: FILE OF INVENTION: Antibodies
: FILE REFERENCE: MSK.P-013-2
: CURRENT APPLICATION NUMBER: US/10/075,947A
: CURRENT FILING DATE: 2002-02-13
: PRIOR APPLICATION NUMBER: 09/142,974
: PRIOR FILING DATE: 1998-09-18
: PRIOR APPLICATION NUMBER: PCT/US97/04427

```

```

1  PRIOR FILING DATE: 1997-03-20
2  PRIOR APPLICATION NUMBER: 60/013,703
3  PRIOR FILING DATE: 1996-03-20
4  NUMBER OF SEQ ID NOS: 5
5  SOFTWARE: PatentIn Ver. 2.1
6  SEQ ID NO 3
7  LENGTH: 1176
8  TYPE: DNA
9  ORGANISM: Murine
10 OTHER INFORMATION: 5fl-secFv-streptavidin
11 FEATURES:
12 NAME/KEY: unsure
13 LOCATION: (37)
14 FEATURE:
15 NAME/KEY: unsure
16 LOCATION: (79)
17 OS:10-075-947A-3

```

|                           |        |                   |           |              |
|---------------------------|--------|-------------------|-----------|--------------|
| Query Match               | 46.6%; | Score 443.6;      | DB 15;    | Length 1176; |
| Best Local Similarity     | 76.6%; | Pred. No. 6e-126; |           |              |
| Matches 558; Conservative | 0;     | Mismatches 161;   | Indels 9; | Gaps 1;      |

|    |     |  |     |
|----|-----|--|-----|
| Qy | 8   | AGGTCAAGCTGCAGAGAGTCAAGGGAGAGCTTAGTGTACGCTGAGAGGGTCCGGAAACCTCT   | 67  |
| Db | 2   | AGGTGAAATCTGCAGCAGTCAAGGACCTGAACTGTGTAGACCTGTGGGCTTCACTGAAGATAT  | 61  |
| Qy | 68  | CTGTGTACACCTCTGTGATTCACCTTTCAGTAGCTTTGAAATGCATCTGGGTGTGTAGAGCTC  | 127 |
| Db | 62  | CTGCAAGACTTCTGGANACAATTCATCTGAATACACCATGCACTGGGGTGAAGACAGAGCC    | 123 |
| Qy | 128 | CAGAGAAAGGGCTCGAGTGGGTTCGATATATTAGTAGTGGCAGTAGTACATCTATATG       | 187 |
| Db | 122 | ATGGAAGAAGCCTTAGTGTGATTTGAAGGATTAATTCCTAAACAATGGTGTACTTACTACA    | 183 |
| Qy | 188 | CAGACACAGTGAAGGGACGATTCACCATCTCCAGAGACAAATCCCAAGAACCCCTGTTC      | 247 |
| Db | 182 | AGCAGAAAGTTCAAGGGCAAGGGCCATTTAGTCTGTAGACAAAGTCTCCGACAGACAGCTACA  | 243 |
| Qy | 248 | TGCAAAATGACACAGTCTTAGGTCTGAGACACGGTCACTGATTTACTGTGCAGAGAGTTACG   | 307 |
| Db | 242 | TGGAGTCCGCAAGCCTGACATCTGAGGATTCGTGCAGTCTATTACTGTGCAGAGACTACTA    | 303 |
| Qy | 308 | GG-----GCTTATTGGGGGCCAAAGGGACCAAGGTCACGGTCAAGCTTCCTCAGGTGAGGCG   | 358 |
| Db | 302 | CGGTCCCGTTTGCTTACTGGGTGTCAAAGGACCAAGGTCAACGGTCTCTCAAGTGAAGCG     | 364 |
| Qy | 359 | GCTCAGGCGGAGGTGTCTCTGCGGTGGCGGATCGGACATTGAGTCAACCAAGTCTCCAG      | 418 |
| Db | 362 | GTTCAAGCGGAGGTGTCTCTGCGGTGGCGGATCGGACATGAGTCACTCAGTCTCCAG        | 422 |
| Qy | 419 | CAATCATGTCTGCATCTCCAGGGGAGAGGGTCACTACATGACTCGCAGTGGCCAGTTCAAGTG  | 477 |
| Db | 422 | CAATCATGTCTGCATCTCCAGGGGAGAGAGTCACTACATGACTCGCAGTGGCCAGTTCAAGTA  | 481 |
| Qy | 479 | TAAGTATACATGAACTGTGTTCCAAACGAAGTCAGGCACTCCGCCAAAGATGATATTATG     | 538 |
| Db | 482 | TAAGTATACATGATGTATACAGCAGAAAGCTGTCACTCCGCCAAAGATGATATTATG        | 542 |
| Qy | 539 | ACAACATCCAACTGTCTTGTGAGTCCCTGTCTCGTTCAGTGGCAGTGGGTCTGGAGACT      | 598 |
| Db | 542 | ACAACATCCAACTGTGCTTGTGAGTCCCTGTCTCGTTCAGTGGCAGTGGGTCTGGAGACT     | 602 |
| Qy | 599 | CTTACTCTCTCAATCAGACAGACATGAGAGGCTGAAAGTGTGCCACTTATCACTGCCAGC     | 658 |
| Db | 602 | CTTATTTCTCTCAACATCAGACAGACATGAGAGGCTGATGATGTGCCACTTATTTACTGCCATC | 662 |
| Qy | 659 | AGTGAAGTAGTAACCACTCAAGTTCGGTGTGGAGCCAAAGCTGAGAGCTGAAACGGGCGG     | 718 |
| Db | 662 | AGCGAGTAGTATACCGGCTCACTTTCGGTGTGGAGCACAGTTGGAATTAATAACGGGCGG     | 722 |
| Qy | 719 | CCGCGAAT 726   |     |

Db 722 CCGCTGA 729

RESULT 8  
US-09-808-037-5  
; Sequence 5, Application US/09808037  
; Patent No. US20020052311A1

GENERAL INFORMATION:  
; APPLICANT: SOLOMON, Beke  
; APPLICANT: HANAN, Eliat  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF  
; TITLE OF INVENTION: NEUROLOGICAL DISEASES AND DISORDERS  
; FILE REFERENCE: SOLOMON-2D

CURRENT APPLICATION NUMBER: US/09/808,037  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 09/629,971  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 09/473,653  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: US 60/152,417  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.0

SEQ ID NO 5  
; LENGTH: 717  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(717)  
US-09-808-037-5

Query Match 46.5%; Score 442; DB 9; Length 717;  
Best Local Similarity 77.4%; Pred. No. 1,6e-125;  
Matches 554; Conservative 0; Mismatches 150; Indels 12; Gaps 1;

QY 8 AGGTCAAGTGCAGAGAGTGAAGGGGAGGCTAGTGAAGCTGAGAGGCTCCGGAACCT 67  
DB 2 AGGTCAAACTGCAGAGAGTGAAGGGGAGGCTAGTGAAGCTGAGAGGCTCCGGAACCT 61  
QY 68 CCGTGCAGAGCTCTGAGATTCATCTTCAAGAGCTTGGAGAGTGAAGCTGAGAGGCTC 127  
DB 62 CCGTGCAGAGGCTCTGAGATTCATCTTCAAGAGCTTGGAGAGTGAAGCTGAGAGGCTC 121  
QY 128 CAGAGAGAGGCTGAGAGTGAAGGGGAGGCTAGTGAAGCTGAGAGGCTGAGAGGCTC 187  
DB 122 ATGCAAGAGTCTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 181  
QY 188 CAGAGAGTGAAGAGGAGATTCACATCTCCAGAGAGCAATCCCAAGAACACCTGTTCC 247  
DB 182 ACCAGAGATTCAAGGGCAAGGCCACAAATGATGTAGAAATCTCCAGAGAGCTTATA 241  
QY 248 TSCAATGACAGCTCTAAGTCTGAGAGACAGGTCATGATTAATCTGAGAGAG----- 301  
DB 242 TGGAACTGCCAGATGACATCTGAGAGATTCGCCATCTTAATCTGTGAGAGAGGAGCTA 301  
QY 302 -----ATTACGGGGCTTAATTGGGGCCAAAGGACCAAGGTCACCGTCTCTCAGGTGAG 355  
DB 302 CTATGCTCTACTTTGACATCTGAGGAGCAAGTACCAAGTACCGGTCACCGTCTCTCAGGTGAG 361  
QY 356 GGGGCTCAGGAGGAGTGGCTCTGGGGGAGGAGTGGAGTGGAGCATTTGAGAGTCAAGGCTC 415  
DB 362 GGGGCTCAGGAGGAGTGGCTCTGGGGGAGGAGTGGAGTGGAGCATTTGAGAGTCAAGGCTC 421  
QY 416 CAGCAATCATGCTGATCTCAGAGGAGAGAGGTCACATGACCTGCAAGTGAAGTGAAGTGAAG 475  
DB 422 CAGCAATCATGCTGATCTCAGAGGAGAGAGGTCACATGACCTGCAAGTGAAGTGAAGTGAAG 481  
QY 476 GTGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 535  
DB 482 GTGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 541  
QY 536 ATGACATCATCCAAACTGCTCTTGAAGTCCGCTGCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 595

Db 542 ATGACATCATCCAAACTGCTCTTGAAGTCCGCTGCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 601  
QY 536 CCGTCTACTCTCAGATTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 655  
DB 602 CCGTCTACTCTCAGATTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 661  
QY 656 AGCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 711  
DB 662 ATGACGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 717

RESULT 9  
US-10-162-889-5  
; Sequence 5, Application US/10162889  
; Publication No. US20030077252A1  
; GENERAL INFORMATION:  
; APPLICANT: SOLOMON, Beke  
; APPLICANT: HANAN, Eliat  
; TITLE OF INVENTION: AGENTS AND COMPOSITIONS AND METHODS UTILIZING SAME  
; TITLE OF INVENTION: USEFUL IN DIAGNOSING  
; TITLE OF INVENTION: AND/OR TREATING OR PREVENTING PLAQUE FORMING DISEASES  
; FILE REFERENCE: SOLOMON-2B

CURRENT APPLICATION NUMBER: US/10/162,889  
; CURRENT FILING DATE: 2002-06-06  
; PRIOR APPLICATION NUMBER: US/09/629,971  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 09/473,653  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: US 60/152,417  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 717  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(717)  
US-10-162-889-5

Query Match 46.5%; Score 442; DB 14; Length 717;  
Best Local Similarity 77.4%; Pred. No. 1,6e-125;  
Matches 554; Conservative 0; Mismatches 150; Indels 12; Gaps 1;

QY 8 AGGTCAAGTGCAGAGAGTGAAGGGGAGGCTAGTGAAGCTGAGAGGCTCCGGAACCT 67  
DB 2 AGGTCAAACTGCAGAGAGTGAAGGGGAGGCTAGTGAAGCTGAGAGGCTCCGGAACCT 61  
QY 68 CCGTGCAGAGCTCTGAGATTCATCTTCAAGAGCTTGGAGAGTGAAGCTGAGAGGCTC 127  
DB 62 CCGTGCAGAGGCTCTGAGATTCATCTTCAAGAGCTTGGAGAGTGAAGCTGAGAGGCTC 121  
QY 128 CAGAGAGAGGCTGAGAGTGAAGGGGAGGCTAGTGAAGCTGAGAGGCTGAGAGGCTC 187  
DB 122 ATGCAAGAGTCTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 181  
QY 188 CAGAGAGTGAAGAGGAGATTCACATCTCCAGAGAGCAATCCCAAGAACACCTGTTCC 247  
DB 182 ACCAGAGATTCAAGGGCAAGGCCACAAATGATGTAGAAATCTCCAGAGAGCTTATA 241  
QY 248 TSCAATGACAGTCTAAGTCTGAGAGACAGGTCATGATTAATCTGAGAGAG----- 301  
DB 242 TGGAACTGCCAGATGACATCTGAGAGATTCGCCATCTTAATCTGTGAGAGAGGAGCTA 301  
QY 302 -----ATTACGGGGCTTAATTGGGGCCAAAGGACCAAGGTCACCGTCTCTCAGGTGAG 355  
DB 302 CTATGCTCTACTTTGACATCTGAGGAGCAAGTACCAAGTACCGGTCACCGTCTCTCAGGTGAG 361  
QY 356 GGGGCTCAGGAGGAGTGGCTCTGGGGGAGGAGTGGAGTGGAGCATTTGAGAGTCAAGGCTC 415  
DB 362 GGGGCTCAGGAGGAGTGGCTCTGGGGGAGGAGTGGAGTGGAGCATTTGAGAGTCAAGGCTC 421





|    |     |   |     |
|----|-----|---|-----|
| QY | 68  | CTCTGACGCTCTGGAATTCATTACTTTCACTAGCTTTGGAATGCACTGAGGTTGTCAGAGCTC   | 127 |
| Db | 62  | CTCTCAAGGGTCTTGCTACACATTCACTAATATATGCTATGCACTGGGTGAAGCAGATC       | 121 |
| QY | 128 | CAGGAAGGGGCTGGAGTGGGTCGATATATATTAAGTAGGGCAGTAGATCAATCTATATAG      | 187 |
| Db | 122 | ATGCAAAAGAGCTTAGAGTGGATGGATGGATTATTAAGTACTTAATATGATGATCTAGTACA    | 181 |
| QY | 188 | CAGACACAGTAAAGGACGAATTCAACATCTCCAGAGCAATCCCAAGAAACCCCTGTCC        | 247 |
| Db | 182 | ACCGAAGATTCAAGGGCAAGGCCACAAATGACTGTATGAAATATCTCCAGCAAGCTTATA      | 241 |
| QY | 248 | TGCAATATGACCAAGTCTAAGGCTTGAGGACACGGTCACTGATATTAATTCTGTCAAGAG----- | 301 |
| Db | 242 | TGAACTTGCCAGACTGACATCTGAGATTTCTGCATCTATTAATCTGTCAAGAGGGCTA        | 301 |
| QY | 302 | -----ATTACGGGGCTTATTTGGGGCCAAAGGACACGGTCAACGGTCTCTCTAGGTGAG       | 355 |
| Db | 302 | CTATGTCCTAATTTTGACTACTAGGGGCCAAAGTGAACACCGTCAACCGTCTCTCTAGGTGAG   | 361 |
| QY | 356 | GGGGTCAGGGCGAGTGGTCTCTGGCGGTGGGGGATTCGGAATTAGAGTCAACCAAGTTC       | 415 |
| Db | 362 | GGGTTCAAGGGGAGTTGGCTCTGGCGGTGGGGGATTCGGAATTAGAGTCACTCACTCATCTC    | 421 |
| QY | 416 | CAGCAATCATGTCGATCTCCAGGGGAGAGGGTCAACCATGACCTGCAAGTGCACATTCAA      | 475 |
| Db | 422 | CAGCAATCATGTCGATCTCCAGGGGAGAGGATCACCATGACCTGCAAGTGCACAGCTCAA      | 481 |
| QY | 476 | GTTAAGGTAATGAATCTGTTCCAAAGAAAGTCAAGGACCTCCCCCAAAAGATGATTT         | 535 |
| Db | 482 | GTAATAGGTAATGATCACTGTTATCAGCAGAAAGCAGGGACCTCCCCCAAAAGATGATTT      | 541 |
| QY | 536 | ATGACACATCCAAACTGTCTCTGGAATCCCTGCTCGTCTCAGTAGGACAGTGGTCTGAGA      | 595 |
| Db | 542 | ATGACACATCCAAACTGGCTCTGGAATCCCTGCTCGTCTCAGTAGGACAGTGGTCTGAGA      | 601 |
| QY | 596 | CTCTTAACTCTCAACAATAGCAGCATGAGGCTGAAGATGCTGCCACTTAATCTAGCC         | 655 |
| Db | 602 | CTCTTAACTCTCAACAATAGCAGCATGAGGCTGAAGATGCTGCCACTTAATCTAGCC         | 661 |
| QY | 656 | AGCAGTGAAGTAGTAAACCACTCAAGTTCGGTCTGAGGACCAAGTGAAGCTGAAA           | 711 |
| Db | 662 | ATTCAGCGAGTAGTAAACCACTTCAAGTTCGGAGGGGGGACCAAGTGAAGAAATTTAAA       | 717 |





FILE REFERENCE: MSK-P-035-US  
CURRENT APPLICATION NUMBER: US/08/940,544B  
CURRENT FILING DATE: 1997-09-30  
EARLIER APPLICATION NUMBER: PCT/US97/04427  
EARLIER FILING DATE: 1997-03-20  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 717  
TYPE: DNA  
ORGANISM: HUMAN  
FEATURE:  
OTHER INFORMATION: 5F11-bcfv  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (37)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (79)  
US-08-940-544-3

Query Match 45.7%; Score 434.8; DB 8; Length 717;  
Best Local Similarity 76.5%; Pred. No. 2,6e-123;  
Matches 548; Conservative 0; Mismatches 159; Indels 9; Gaps 1;

QY 8 AGGTCAAGCTGCAAGAGTCAAGGGGAGCTTAGTGCAGCTGAGGGGTCCGGAACTCT 67  
DB 2 AGGTGAACTGCAGCAGCTGAGCACTGAACTGGTGNAGCTGGGGCTTCAGTGAAGATAT 61  
QY 68 CCTGTCAGAGCTCTGTGATTCCTTTGAGTTCAGTTCGAGTTCGAGGGCTC 127  
DB 62 CCTGCAAGCTTCTGGAACAATTCAGTAATACACATGCACTGGGTGAAGAGGCC 121  
QY 128 CAGAGAGGGGCTGAGTGGTTCGATATATTAGTGGCAGTAGTACATCTACTATG 187  
DB 122 ATGGAAGAAGCCTTAGTGGATTGAGGATTTATCTTAACATGGTGTACTAATACA 181  
QY 188 CAGACACAGTGAAGGACGATTCATCTCCAGAGACAATCCAGAAACCCCTGTTCC 247  
DB 182 AGCAGAAATTCAAGGGCAAGGCCACATGTGATGACAAAGTCTCCAGACAGCCTACA 241  
QY 248 TGCATAATGCCAGTCTAAGSTCTGAGGACACGGTCAATGATTAATCTGCAAGAGATTACG 307  
DB 242 TGGAGCTCCGACCTGACATCTGAGGATTTGCACTGATTAATCTGTGCAAGAGATACTA 301  
QY 308 GG-----GCTTATTGGGGCCAAAGGACCAAGGTCAAGCTCTCTCAGGTGAGGCG 358  
DB 302 CGGTCCCGTTTGTCTTACTGGGTCCAAAGGACCAAGGTCAAGCTCTCTCAGGTGAGGCG 361  
QY 359 GCTCAGGCGGAGGTGCTCTGGGGTGGCGGATTCGACATTGAGCTCACCAAGTCTCCAG 418  
DB 362 GTTCAGGCGGAGGTGCTCTGGGGTGGCGGATTCGACATTGAGCTCACCAAGTCTCCAG 421  
QY 419 CAATCATGCTGATCTCCAGGGGAGAGGTCACCAATGACCTGCAAGTTCGAAGTG 478  
DB 422 CAATCATGCTGATCTCCAGGGGAGAGGTCACCAATGACCTGCAAGTTCGAAGTG 481  
QY 479 TAAGGTACATGAAGTGTTCACACAGAACTCAGGCACTCCCAAAAGATGATTTATG 538  
DB 482 TAAAGTTACATGCACTGTGACAGCAAGACCTGTCACTCCCAAAAGATGATTTATG 541  
QY 539 ACAACATCCAACTGTCTTCTGAGATCCCTGCTGCTTCAAGTGGCAAGTGGGTCTGGAGCT 598  
DB 542 ACAACATCCAACTGTGCTTCTGAGATCCCTGCTGCTTCAAGTGGCAAGTGGGTCTGGAGCT 601  
QY 599 CTTACTCTCTCAATCAGACAGATGAGGCTGAGAGATGCTGCACTTACTACTGACAGC 658  
DB 602 CTTATTCTCTCAACACAGAGATGAGGCTGTAGATGCTGCACTTATTTACTGCAATC 661  
QY 659 AGTGAAGTAGTAACCACTCAGCTTCGCTGGGACCAAGCTGAGCTGAACGG 714  
DB 662 AGCGAGTAGTAAACCGCTCAGCTTCGCTGGGACACAGTTGGAATAAAACGG 717

Search completed: June 23, 2005, 12:47:09  
Job time : 691 secs



Db 499 GATTACGGGGCTTATTTGGGGCCAGAGGACACAGGTCAACGTCCTCTCCAGGTGAGGCGGT 558  
Qy 361 TCAGCGGAGAGTGGCTCTGGCGGTGGCGGATCGGACATTTGAGCTCAACCGAGTCTCCAGCA 420  
Db 559 TCAGCGGAGAGTGGCTCTGGCGGTGGCGGATCGGACATTTGAGCTCAACCGAGTCTCCAGCA 618  
Qy 421 ATCATGTCTGCATCTTCAGAGGAGAGGATCACCATGACCTGCAAGTCCAGTTCAAGTGA 480  
Db 619 ATCATGTCTGCATCTTCAGAGGAGAGGATCACCATGACCTGCAAGTCCAGTTCAAGTGA 678  
Qy 481 AGGTACATGAACCTGGTTCCAGAGAGTACGACCTCCCCCAAAAGATGATTTATGAC 540  
Db 679 AGGTACATGAACCTGGTTCCAGAGAGTACGACCTCCCCCAAAAGATGATTTATGAC 738  
Qy 541 ACATCCAAACTGTCTTCTGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGAGACCTCT 600  
Db 739 ACATCCAAACTGTCTTCTGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGAGACCTCT 798  
Qy 601 TACTCTTCACATCAGCAGCAGATGAGAGGCTGAAGATGCTGCCATTACTGCGCAGCAG 660  
Db 799 TACTCTTCACATCAGCAGCAGATGAGAGGCTGAAGATGCTGCCATTACTGCGCAGCAG 858  
Qy 661 TGGAGTAGTAACCCCACTCACGTTCCGTGCTGGGACCAACCTGGAAGTGAACGGGCGGCG 720  
Db 859 TGGAGTAGTAACCCCACTCACGTTCCGTGCTGGGACCAACCTGGAAGTGAACGGGCGGCG 918  
Qy 721 GCAGAACAAAACATCATCTCAGAGAGATCTGAATGGGGCGGTGACGAC 772  
Db 919 GCAGAACAAAACATCATCTCAGAGAGATCTGAATGGGGCGGTGACGAC 970

## RESULT 2

US-09-618-869-7  
; Sequence 7, Application US/09618869  
; Patent No. 6455279  
; GENERAL INFORMATION:  
; APPLICANT: Ambrosius, Dorthée  
; APPLICANT: Rudolph, Rainer  
; APPLICANT: Schaeffner, Joerg  
; APPLICANT: Schwarz, Elisabeth  
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND  
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR  
; TITLE OF INVENTION: CHAPERONES  
; FILE REFERENCE: 20381  
; CURRENT APPLICATION NUMBER: US/09/618,869  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: EP99114811.5  
; PRIOR FILING DATE: 1999-07-29  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1256  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (199)..(969)  
US-09-618-869-7

Query Match 78.8%; Score 749.6; DB 3; Length 1256;  
Best Local Similarity 98.2%; Pred. No. 2,9e-205;  
Matches 758; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ATGGCGGAGGTCAAGTGCAGAGTCAAGGAGGCTTGTGCAGGCTCGAGAGGTCGCGG 60  
Db 199 ATGGCGGAGGTCAAGTGCAGAGTGTGGGAGGCTTGTGCAGGCTCGAGAGGTCGCGG 258  
Qy 61 AAATCTCTCTGCAAGCTCTGATTCATTTCAAGTCTTGAATGATGATGATGATGAT 120  
Db 259 AAATCTCTCTGCAAGCTCTGATTCATTTCAAGTCTTGAATGATGATGATGATGAT 318  
Qy 121 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCATATATTAGTAGGCGAGTACCATC 180

Db 319 CAGGCTCCAGAGAGGGGCTGAGTGGGTGCATATATTAGTAGGCGAGTACCATC 378  
Qy 181 TACTATGAGACACAGTGAAGGAGCATTTCAACATCTCCAGAGACATTCACCAAGAACCC 240  
Db 379 TACTATGAGACACAGTGAAGGAGCATTTCAACATCTCCAGAGACATTCACCAAGAACCC 438  
Qy 241 CTGTTCTGCAAAATGACCATGTTAAGTCTGAGGACACGGTCTATTTACTGTGCAAGA 300  
Db 439 CTGTTCTGCAAAATGACCATGTTAAGTCTGAGGACACGGTCTATTTACTGTGCAAGA 498  
Qy 301 GATTACGGGGCTTATTTGGGGCCCAAGGACCAACGTCACCTCTCTCAGTGGAGGCGCG 360  
Db 499 GATTACGGGGCTTATTTGGGGCCCAAGGACCAACGTCACCTCTCTCAGTGGAGGCGGT 558  
Qy 361 TCAGCGGAGAGTGGCTCTGGCGGTGGCGGATCGACATTTGAGCTCAACCAAGTCTCCAGCA 420  
Db 559 TCAGCGGAGAGTGGCTCTGGCGGTGGCGGATCGACATTTGAGCTCAACCAAGTCTCCAGCA 618  
Qy 421 ATCATGTCTGCATCTTCAGAGGAGAGGATCACCATGACCTGCAAGTCCAGTTCAAGTGA 480  
Db 619 ATCATGTCTGCATCTTCAGAGGAGAGGATCACCATGACCTGCAAGTCCAGTTCAAGTGA 678  
Qy 481 AGGTACATGAACCTGGTTCCAGAGAGTACGACCTCCCCCAAAAGATGATTTATGAC 540  
Db 679 AGGTACATGAACCTGGTTCCAGAGAGTACGACCTCCCCCAAAAGATGATTTATGAC 738  
Qy 541 ACATCCAAACTGTCTTCTGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGAGACCTCT 600  
Db 739 ACATCCAAACTGTCTTCTGAGATCCCTGCTGCTTCAAGTGGAGTGGGTCTGAGACCTCT 798  
Qy 601 TACTCTTCACATCAGCAGCAGATGAGAGGCTGAAGTGTGCCATTACTGCGCAGCAG 660  
Db 799 TACTCTTCACATCAGCAGCAGATGAGAGGCTGAAGTGTGCCATTACTGCGCAGCAG 858  
Qy 661 TGGAGTAGTAACCCCACTCACGTTCCGTGCTGGGACCAACCTGGAAGTGAACGGGCGGCG 720  
Db 859 TGGAGTAGTAACCCCACTCACGTTCCGTGCTGGGACCAACCTGGAAGTGAACGGGCGGCG 918  
Qy 721 GCAGAACAAAACATCATCTCAGAGAGATCTGAATGGGGCGGTGACGAC 772  
Db 919 GCAGAACAAAACATCATCTCAGAGAGATCTGAATGGGGCGGTGACGAC 970

## RESULT 3

US-08-956-047-24  
; Sequence 24, Application US/08956047  
; Patent No. 5882924  
; GENERAL INFORMATION:  
; APPLICANT: Fritz, Hans-Joachim  
; APPLICANT: Hennecke, Frank  
; APPLICANT: Kolmar, Harald  
; TITLE OF INVENTION: Genetic Selection, by Means of Signal  
; TITLE OF INVENTION: Transduction in Microorganisms, of Proteins which are  
; TITLE OF INVENTION: Capable of Ligand Binding  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flinnegan, Henderson, Farbow, Garrett & Dunner  
; STREET: 1300 I Street, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.

; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,047  
; FILING DATE: 22-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:



```

Db      191 GGAGACAGGTAAAGGGCCGATTCACATCTCCAGAGCAATCCCAAGAACCCCTGTTTC
Oy      247 CTGCAAAATGACCAAGTCTAAGTCTGAGGACACGGTCAATATTACTGCAAGATTAAC
Db      241 CTGCAAAATGACCAAGTCTAAGTCTGAGGACACGGCCTATTAATCTGTCAGAGATTAC
Oy      307 GGGGCTTATTTGGGCGCAAGGAGCAACGCTCCTCTCCAGTGGAGCGGCTCAAGC
Db      301 GGGGCTTATTTGGGCGCAAGGAGCACTGCTCCTCTGCAAGTGGTGGTAGCGGT
Oy      367 GGAGGTGCTCTGGCGGTGGCGGATTCGACATTAAGCTCAACGCTCCACATTAAC
Db      361 GGTCGGGCAAGTGGCGCGCGGCTCTCAATTTGTTCTAACCCAGTCTCCACATTAAC
Oy      427 TCTGCATCTCCAGGGGAGAGGGTCAACATGACCTGCAAGTCCAGTTCAAGTAAAGTAC
Db      421 TCTGCATCTCCAGGGGAGAGGGTCAACATGACCTGCAAGTCCAGTTCAAGTAAAGTAC
Oy      487 ATGAACCTGTTCCAAACAGAGTCAAGGACCTCCCAAAAGATGATTTATGACACATCC
Db      481 ATGAACCTGTTCCAAACAGAGTCAAGGACCTCCCAAAAGATGATTTATGACACATCC
Oy      547 AAACGTCTCTTGGAGTCCCTCTGCTGCTCAAGTGGAGAGGGTCTGGGACCTTTACTCT
Db      541 AAACGTCTCTTGGAGTCCCTCTGCTGCTCAAGTGGAGAGGGTCTGGGACCTTTACTCT
Oy      607 CTCACAAATGACAGCATGAGGCTGAGAGTGCATTAATCACTGACAGAGTGAAGT
Db      601 CTCACAAATGACAGCATGAGGCTGAGAGTGCATTAATCACTGACAGAGTGAAGT
Oy      667 AGTAACCACTCAACGTTGGTGGTGGGACCAAGCTGAGCTGAACG 713
Db      661 AGTAATCCACTCAACGTTGGTGGGACCAAGCTGAGCTGAACG 707

RESULT 5
US-08-190-199A-62
; Sequence 62, Application US/08190199A
; Patent No. 5830663
; GENERAL INFORMATION:
; APPLICANT: EMBLETON, Michael J.
; APPLICANT: GOROCHOV, Guy
; APPLICANT: JONES, Peter T.
; APPLICANT: WINTER, Gregory P.
; TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,199A
; FILING DATE: 13-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01483
; FILING DATE: 10-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9212419.7
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9117352.6
; FILING DATE: 10-AUG-1991
; INFORMATION FOR SEQ ID NO: 62:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 672 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..669
; US-08-190-199A-62

Query Match      61.1%; Score 580.6; DB 2; Length 672;
Best Local Similarity 91.1%; Pred. No. 8.3e-157;
Matches 644; Conservative 0; Mismatches 24; Indels 39; Gaps 1;

Oy      7 GAGGTCAAGCTGACAGGATTCAGGGGAGGCTTAAGTCAGAGCCTTGAGGGGTCCCGAAACTC
Db      1 GATGTGACAGCTGTGGAGTCTGGGGGAGGCTTAAGTCAGAGCCTTGAGGGGTCCCGAAACTC
Oy      67 TCTGTGACAGCTCTGATTAATCACTTTCAGTAGCTTTGGAATGACATGGGTTGTCAGGCT
Db      61 TCTGTGACAGCTCTGATTAATCACTTTCAGTAGCTTTGGAATGACATGGGTTGTCAGGCT
Oy      127 CCAGAGAAAGGGGCTGAGTGGGTCGATATTAATGATGAGCAGTAGTACCATCTACTAT
Db      121 CCAGAGAAAGGGGCTGAGTGGGTCGATATTAATGATGAGCAGTAGTACCATCTACTAT
Oy      187 GCAGACAGTGAAGAGGACATTAACCATCTCCAGAGCAATCCCAAGAACCTCTGTTTC
Db      181 GCAGACAGTGAAGAGGCGCATTAACCATCTCCAGAGCAATCCCAAGAACCTCTGTTTC
Oy      247 CTGCAAAATGACCAAGTCTAAGTCTGAGGACACGGTCACTGATTAATCTGCAAGATTAAC
Db      241 CTGCAAAATGACCAAGTCTAAGTCTGAGGACACGGCCTATTAATCTGCAAGATTAAC
Oy      307 GGGGCTTATTTGGGCGCAAGGAGCAACGCTCACCGTCTCCAGTGGAGCGGCTCAAGC
Db      301 GGGGCTTATTTGGGCGCAAGGAGCACTGCTCACTGCTCTGCAAGCTAG-----
Oy      367 GGAGTGGCTCTGGCGGTGGCGGATTCGACATTAAGCTCAACCCAGTCTCCACATTAAC
Db      348 -----CCAAATTTGTTCCACCAATCTCCACCAATTAATG 381
Oy      427 TCTGCATCTCCAGGGGAGAGGTCACATGACCTGCAAGTCCAGTTCAAGTAAAGTAC
Db      382 TCTGCATCTCCAGGGGAGAGGTCACATGACCTGCAAGTCCAGTTCAAGTAAAGTAC
Oy      487 ATGAACCTGTTCCAAACAGAGTCAAGGACCTCCCAAAAGATGATTTATGACACATCC
Db      442 ATGAACCTGTTCCAAACAGAGTCAAGGACCTCCCAAAAGATGATTTATGACACATCC
Oy      547 AAACGTCTCTTGGAGTCCCTCTGCTGCTCAAGTGGAGAGGGTCTGGGACCTTTACTCT
Db      502 AAACGTCTCTTGGAGTCCCTCTGCTGCTCAAGTGGAGAGGGTCTGGGACCTTTACTCT
Oy      607 CTCACAAATGACAGCATGAGGCTGAGAGTGCATTAATCACTGACAGAGTGAAGT
Db      562 CTCACAAATGACAGCATGAGGCTGAGAGTGCATTAATCACTGACAGAGTGAAGT
Oy      667 AGTAACCACTCAACGTTGGTGGTGGGACCAAGCTGAGCTGAACG 713
Db      622 AGTAATCCACTCAACGTTGGTGGGACCAAGCTGAGCTGAACG 668

RESULT 6
US-08-956-047-29
; Sequence 29, Application US/08956047
; Patent No. 5882924
; GENERAL INFORMATION:
; APPLICANT: Pritz, Hans-Joachim
; APPLICANT: Hennecke, Frank

```

APPLICANT: Kolmar, Harald  
TITLE OF INVENTION: Genetic Selection, by Means of Signal  
TITLE OF INVENTION: Transduction in Microorganisms, of Proteins which are  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pinegar, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,047  
FILING DATE: 22-OCT-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/257,669  
FILING DATE: 08-JUN-1994  
APPLICATION NUMBER: DE P 43 19 296.3  
FILING DATE: 10-JUN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pohlman, Sandra M.  
REGISTRATION NUMBER: P39,691  
REFERENCE/DOCKET NUMBER: 05552.1368-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ. ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 793 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..354  
OTHER INFORMATION: /note= "Molecules 16-354 encode a  
OTHER INFORMATION: peptide."  
NAME/KEY: CDS  
LOCATION: 383..775  
OTHER INFORMATION: /note= "Molecules 383-775 encode a  
OTHER INFORMATION: peptide."  
US-08-956-047-29

Query Match 57.2%; Score 543.6; DB 2; Length 793;  
Best Local Similarity 85.2%; Pred. No. 3.9e-146;  
Matches 654; Conservative 0; Mismatches 59; Indels 55; Gaps 2;

QY 2 TGGCCAGAGTCAAGTGCAGAGTCAAGGAGGCTTATGTCAGCGCTTGAAGGATCCCGGA 61  
DB 8 TGCATCACTACGTCAGCTGCTGAGACTTGGGGAGGCTTTGTCAGCGCTTGAAGGATCCCGGA 67  
QY 62 AACTCTCTGTGAGGCTGCTGATTCATCTTCACTAGTACCTTGAATGCACTGGGTTGCTC 121  
DB 68 AACTCTCTGTGAGGCTGCTGATTCATCTTCACTAGTACCTTGAATGCACTGGGTTGCTC 127  
QY 122 AGGCTCCAGAGAGGGGCTGAGTGGGTGTCATATATTAAGTAGGTCAGTAGTACATCT 181  
DB 128 AGGCTCCAGAGAGGGGCTGAGTGGGTGTCATATATTAAGTAGGTCAGTAGTACATCT 187  
QY 182 ACTATGCAAGACAGTGAAGGAGCATTCACCTTCCAGAGCANTCCCAAGAACACCC 241  
DB 188 ACTATGCAAGACAGTGAAGGAGCATTCACCTTCCAGAGCANTCCCAAGAACACCC 247

QY 242 TGTCTCTGCAATGACCACTTAAGTCTGAGACACGGTCAATGATTAATCTGTCAGAG 301  
DB 248 TGTCTCTGCAATGACCACTTAAGTCTGAGACACGGTCAATGATTAATCTGTCAGAG 307  
QY 302 ATTACGGGGCTTATTGGGGCCAAAGGACACAGTCCCTCTCTCCTGAGTGGAGGCG 358  
DB 308 ATTACGGGGCTTATTGGGGCCAAAGGACACAGTCCCTCTCTCCTGAGTGGAGT 367  
QY 359 -----GCTCAGGCGAGTGGCT 376  
DB 368 TCATTTGTACATGAGAAATATAAGTAAACAAAGACATATTGACACTGACACTTAACG 427  
QY 377 CTGGCGGTGGCGGATCGAGC-----ATTGAGCTACCCAGTCTCCAGATCATG 426  
DB 428 TTACTGTTTACCCCTGTGTCAAAAGCCCAAAATTTGTTCTACCAAGTCTCCAGAACATG 487  
QY 427 TCGCATCTCCAGGGGAGGGGTCAACATGACCTGAGGCGCAGTTCAAGTGAAGGTAC 486  
DB 488 TTGCATCTCCAGGGGAGGAAAGTCAACATGACCTGAGGCGCAGTTCAAGTGAAGGTAC 547  
QY 487 ATGAACGTGTTCCACAGAGTCAAGGACCTCCCAAAAGATGATTTATGACACATCC 546  
DB 548 ATGAACGTGTTCCACAGAGTCAAGGACCTCCCAAGGCGTGGACATATGACATATC 607  
QY 547 AAACGTCTTTCTGAGTCTCTGCTGCTTCAAGTGGAGTGGTCTTGAACCTTTACTCT 606  
DB 608 AAACGTCTTTCTGAGTCTCTGCTGCTTCAAGTGGAGTGGTCTTGAACCTTTACTCT 667  
QY 607 CTCACATGAGAGCATGAGGCTGAAGATGCTGCACCTTACTACTGACAGCATGAGT 666  
DB 668 CTCACATGAGAGCATGAGGCTGAAGATGCTGCACCTTACTACTGACAGCATGAGT 727  
QY 667 ACTAACCCCACTACGTTGCTGCTGAGCAACAGTGGAGCTGAACCG 714  
DB 728 AGTATCCACTCACTTCCGTGCTGAGCAACAGTGGAGCTGAACCG 775

RESULT 7  
US-08-190-199A-60  
Sequence 60, Application US/08190199A  
Patent No. 5830663  
GENERAL INFORMATION:  
APPLICANT: EMBLETON, Michael J.  
APPLICANT: GOROCHEV, Guy  
APPLICANT: JONES, Peter T.  
APPLICANT: WINTER, Gregory P.  
TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS  
NUMBER OF SEQUENCES: 70  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,199A  
FILING DATE: 13-JUL-1994  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/01483  
FILING DATE: 10-AUG-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: GB 9212419.7  
FILING DATE: 11-JUN-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: GB 9117352.6  
FILING DATE: 10-AUG-1991

```

; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 708 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHEICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..705
; US-08-190-199A-60

Query Match      48.7%; Score 462.8; DB 2; Length 708;
Best Local Similarity 79.5%; Pred. No. 6.3e-123;
Matches 561; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 8 AGGTCAAGCTGACGAGGTGAGGGGAGGCTTGTGACAGCTGAGAGGTCCCGAAACTCT 67
DB 2 AGGTGACGCTGAAGAGGAGTCAAGACCTGGCTGGGCGCTTCAAGAGCTGTCCATCA 61
QY 68 CCTGTGACAGCTCTGGAATTCATTTCAGTGTGGAATGACCTGGGTTCTCAGGCTC 127
DB 62 CTTCGACTGTCTCTGGGTTTTCATTACCAAGCTATGTTACACTGGTTCGACGCTC 121
QY 128 CAGAGAAAGGGGCTGAGTGGGTGCGCATATATTAGTAGTGAGAGTACCATCTATG 187
DB 122 CAGGAAAGGGGTCTGAGTGGGTGAGTATATGAGGCTGTGGAAGCAAAATTATATT 181
QY 188 CAGACACACTGAAGGAGCATTCACATCTCCAGAGCAATCCCAAGAACACCCGTTC 247
DB 182 CGG---CTCTCATGTCCACACTGAGCATGAGAAAGCAACTCCAGAGCAAGTTTCT 238
QY 248 TCGAATGACCAAGCTTAAGTCTAGAGACACGGTCACTATTACTGTGCAAGATTACG 307
DB 239 TAAAATGACAGTCTGCAAACTGATGACACAGCCATGACTACTGTGCAAGATCGGG 298
QY 308 GGGCTTATTTGGGCGCAAGGAGCAAGGTACCGTCTCTCAAGTGGAGGGGCTCAGGCG 367
DB 289 GGGCTTATTTGGGCGCAAGGAGCAAGGTACCGTCTCTCAAGTGGAGGGGCTCAGGCG 358
QY 368 GAGGTGGCTCTGGCGGTGGCGGATCGGACATTGAGTCCCAAGTCTCCAGCAATCATGT 427
DB 359 GTGGCGGAGTGGCGGCGGCTCTCAAAATTGTTCTCAAGGCTCCAGCAATCATGT 418
QY 428 CTGCATCTCCAGGGGAGAGGTCAACATGACCTGACAGTCCAGTTCAGATGTAAGTACA 487
DB 419 CTGCATCTCCAGGGGAGAGGTCAACATGACCTGACAGTCCAGTTCAGATGTAAGTACA 478
QY 488 TGAACCTGTTCCAAACAGAGTCAAGGACCTCCGCCCAAAAGATGATTATGACATCCA 547
DB 479 TGCACCTGGTGAACAGAGTCAAGGACCTCCGCCCAAAAGATGATTATGACATCCA 538
QY 548 AACTGCTCTTCTGAGTCCCTGCTGCTTGAAGTGGAGTGGGTCTGGAACTTTACTCTC 607
DB 539 AACTGCTCTTCTGAGTCCCTGCTGCTTGAAGTGGAGTGGGTCTGGAACTTTACTCTC 598
QY 608 TCACATTCAGCAGCATGAGGCTGAAGATGCTGCCACTTACTACTGCCAGAGTGGAGTA 667
DB 599 TCACATTCAGCAGCATGAGGCTGAAGATGCTGCCACTTACTACTGCCAGAGTGGAGTA 658
QY 668 GTAACCCACTCAAGTTCGGTGTGGAGCCAAAGCTGAGGCTGAAGCG 713
DB 659 GTAACCCACTCAAGTTCGGTGTGGAGCCAAAGCTGAGGCTGAAGCG 704

RESULT 8
US-08-564-164A-1
; Sequence 1, Application US/08564164A
; Patent No. 6159947
; GENERAL INFORMATION:
; APPLICANT: Schweighoffer, Fabien
```

```

; APPLICANT: Tocque, Bruno
; TITLE OF INVENTION: Intracellular Binding Proteins and Use
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Road, 3643
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,164A
; FILING DATE: 28-DEC-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR94/00714
; FILING DATE: 15-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93/07241
; FILING DATE: 16-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin F.
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: ST93030-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3816
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 858 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHEICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..858
;   FEATURE:
;     NAME/KEY: misc feature
;     LOCATION: 442..486
;   OTHER INFORMATION: /product= "Linker"
;   NAME/KEY: misc feature
;     LOCATION: 82..810
;   OTHER INFORMATION: /product= "ScFv anti-Ras"
; US-08-564-164A-1

Query Match      48.6%; Score 462.2; DB 3; Length 858;
Best Local Similarity 77.0%; Pred. No. 1e-122;
Matches 598; Conservative 0; Mismatches 158; Indels 21; Gaps 2;

QY 1 ATGGCCGAGGTCAAGCTGACGAGTCAAGGAGGAGGCTTAGTGCAGGCTGAGAGGCTCCGG 60
DB 76 ATGGCTCAGGTGAACCTGACGAGTCAAGGAGGAGGCTTAGTGCAGGCTGAGAGGCTCCGTG 135
QY 61 AAACCTCTCTGAGGCTCTGATTCATTTCAGTACCTTTGGAATGACATGAGTGGTCTGT 120
DB 136 AAACCTCTCTGATGATGCTCTGATTCATTTCAGTACCTTTGGAATGACATGAGTGGTCTGT 135
QY 121 CAGGCTCAAGAGAGGGGCTGAGTGGGTGCGCATATATTAGTAGTGAGTGAATCAATC 180
DB 196 CAGACTCAGGAGGAGGAGCTGAGTGGGTGCGCATATATTAGTAGTGAGTGAATCAATC 255
QY 181 TACTATGAGACACAGTGAAGGAGCATTCACATCTCCAGAGCAATCCCAAGAACACC 240
```



Db 256 TACTATGAGAAAGGCGATTCACCATCTCCAGAGACAAATGCCAAGAACCC 315  
 Qy 241 CTGTTCTCTGCAATATGACAGATCTTAAGATCTGAGAGACAGGCTCATGTATTACTGTGAGA 300  
 Db 316 CTGTACTCTCAATATGACAGATCTGAGATCTGAGAGACAGTCCCTTGTATTACTGTGAGA 375  
 Qy 301 GATTACG-----GCTTATTGAGGCGCAAGAGAACCAAGCTCACCGTC 342  
 Db 376 CATGAGGAGTACGGGTACCGACTTCTTGTATTACTGAGGCGCAAGAGAACCAAGCTCACCGTC 435  
 Qy 343 TCTCTGAGTGAAGGCGGCTCAGGCGGAGGCTCTGCGGCTGCGGATTCGACATTGAG 402  
 Db 436 TCTCTGAGTGAAGGCGGCTCAGGCGGAGGCTCTGCGGCTGCGGATTCGACATTGAG 495  
 Qy 403 CTCACCACTCTCCAGCAATCATGTCTGATCTCCAGGAGAGAGGCTCATGACCTGC 462  
 Db 496 CTCACCACTCTCCAGCAATCATGTCTGATCTCTGAGAGAACTGTCTTCATCATGATGT 555  
 Qy 463 AGTCCAGT---TCAAGTGAAGTACATGAACCTGCTCCACAGAAAGTCAAGCACTCC 519  
 Db 556 CTAGCAAGTGAAGGCAATTCATTAATTAGCGTGTATCAGCAAGCAAGGGAATCT 615  
 Qy 520 CCGCAAAAGATGATTTATGACATCCAACTGTCTTGTGAGTCCCTGCTGCTTCACT 579  
 Db 616 CCTCAGCTCTGATCTATTATGCAAGTCTGAGATGAGAGATGAGGCTCCATCAAGCTTCACT 675  
 Qy 580 GGCAGTGGGTCTGAGGAGCTTCTTCTCTCACAATCAGCAGCATGAGAGGCTGAAGTCT 639  
 Db 676 GGCAGTGGATCTGAGCAGAGTTTCTCTCAGATCAGCAACATGCAACTGAAATGAA 735  
 Qy 640 GGCATTTACTACTGCGAGAGTGAAGTAAACCACTCACTGCTGCTGAGACCAAG 699  
 Db 736 GGGGTTTATTACTGTCAACAGGCTTCAAGATCTTCCAGCTTGGAGCTGACCAAG 795  
 Qy 700 CTGAGACTGAAGGCGGCGCGGCAAACTCATCTCAGAAAGATCTGAT 756  
 Db 796 CTGGAATTAACAGGCGCGCGGCAAACTCATCTCAGAAAGATCTGAT 852

## RESULT 9

US-09-674-677-4  
 / Sequence 4, Application US/09674677  
 / Patent No. 6562622  
 / GENERAL INFORMATION:  
 / APPLICANT: Cola, et al.  
 / TITLE OF INVENTION: CONTINUOUS IN-VITRO EVOLUTION  
 / FILE REFERENCE: 674537-2003  
 / CURRENT APPLICATION NUMBER: US/09/674, 677  
 / CURRENT FILING DATE: 2000-12-11  
 / PRIOR APPLICATION NUMBER: PCT/AU99/00341  
 / PRIOR FILING DATE: 1999-05-07  
 / PRIOR APPLICATION NUMBER: AU PP3445  
 / PRIOR FILING DATE: 1998-05-08  
 / NUMBER OF SEQ ID NOS: 38  
 / SOFTWARE: PatentIn version 3.0  
 / SEQ ID NO 4  
 / LENGTH: 807  
 / TYPE: DNA  
 / ORGANISM: Homo Sapiens  
 / FEATURE:  
 / NAME/KEY: misc.feature  
 / LOCATION: (1)-(807)  
 / OTHER INFORMATION: Sequence of the anti-hepatitis surface antigen (4c2) scfv  
 / Patent No. 6562622  
 US-09-674-677-4

Query Match 48.4%; Score 460.2; DB 4; Length 807;  
 Best Local Similarity 77.5%; Pred. No. 3, 7e-122;  
 Matches 574; Conservative 0; Mismatches 158; Indels 9; Gaps 1;  
 Qy 1 ATGGCCGAGTCAAGTCAAGAGTCAAGGAGGCTTACTGACGCTGAGGCTCCCG 60  
 Db 3 ATGGCCGAGTCAAGTCAAGAGTCAAGGAGGCTTACTGACGCTGAGGCTCCCG 62

Qy 61 AAATCTCTGTCAGACTCTGATTCACATTACAGTACGTTGAAATGACATGAGTTCGT 120  
 Db 63 AGATTAATCTGCAAGGCTTCCGCTACATTCACATGATTAATGCTATGATGAGTGGAG 122  
 Qy 121 CAGGCTCCAGAGAGGCGCTGAGTGGTTCGATTAATTAATTAATGAGTACATTC 180  
 Db 123 CAGAGTCATGCCAAGAGTCTAGAGTGAATGAGCTTAATTAATTCCTTGTGTAATCA 182  
 Qy 181 TACTATGACAGACAGTGAAGGAGATTCACATCTCCAGAGCAATCCCAAGAACCC 240  
 Db 183 AACTACACCGAAGTTTGAAGCCAGGCAATATCTGTAGCAAAATCTTCAACACA 242  
 Qy 241 CTGTTCTCTGCAATATGACAGTCTTAAGTCTGAGAGACGCTCATGTTATTAAGTGAAGA 300  
 Db 243 GGTATTGGAATCTGAGATTTGATCATGAGAGATTCGCAATTAATTAAGTGAAGA 302  
 Qy 301 G-----ATTACGGGCTTATTGGGCAAGGACCAAGTCAAGCTCTCTCAAGT 351  
 Db 303 GTGATGACTGTGCTTTCATGTCTGAGGCAAGGACCAAGCTCAAGCTCTCTCAAGT 362  
 Qy 352 GAGAGGCTCAAGGAGGAGTGTGCTGAGGAGTGGAGATCGGACATTAAGTCAAGC 411  
 Db 363 GAGAGGCTTCAAGGAGGAGTGTGCTGAGGAGTGGAGATCGGACATTAAGTCAAGC 422  
 Qy 412 TCTCCAGATCATGTCTGATCTCCAGGAGAGAGGCTCAACATGACCTGAGTGCAGT 471  
 Db 423 TCTCCAGATCATGTCTGATCTCCAGGAGAGAGTCAACATGACCTGAGTGCAGC 482  
 Qy 472 TCAAGTGAAGTCAATGAATCTGCTTCAACAGAGTCAAGCACTTCCCAAAAGATG 531  
 Db 483 TCAAGTGAAGTCAATGAATCTGCTTCAACAGAGTCAAGCACTTCCCAAAAGATG 542  
 Qy 532 ATTATGACATCAACATGCTTCTGAGAGTCCGCTGCTTCAAGTGGAGTGGTCT 591  
 Db 543 ATTATGACATCAACATGCTTCTGAGAGTCCGCTGCTTCAAGTGGAGTGGTCT 602  
 Qy 592 GGGACCTTACTCTCTCAATCAATCAGACAGATGAGGCTGAGATCTGCACTTCACTAC 651  
 Db 603 GGGACCTTACTCTCTCAATCAATCAGACAGATGAGGCTGAGATCTGCACTTCACTAC 662  
 Qy 652 TGCACAGTGAAGTGAATCAACCTCAGCTTGTGCTGAGAGCAAGCTGAGAGTGA 711  
 Db 663 TGCACAGTGAAGTGAATCAACCTCAGCTTGTGCTGAGAGCAAGCTGAGAGTGA 722  
 Qy 712 CGGCGCGCGCGAGCAAA 732  
 Db 723 CGGCGCGCGCGAGATTATAA 743

## RESULT 10

US-08-661-052-15  
 / Sequence 15, Application US/08661052  
 / Patent No. 5837243  
 / GENERAL INFORMATION:  
 / APPLICANT: Yashwant M. Deo  
 / APPLICANT: Joel Goldstein  
 / APPLICANT: Robert Graziano  
 / APPLICANT: Chezia Somasundaram  
 / TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISD  
 / NUMBER OF SEQUENCES: 16  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: LAHYE & COCKFIELD  
 / STREET: 60 State Street, Suite 510  
 / CITY: Boston  
 / STATE: Massachusetts  
 / COUNTRY: USA  
 / ZIP: 02109-1875

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/661,052  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/484,172  
FILING DATE: 07-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MXI-043CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1679 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 11..1667  
US-08-661-052-15

Query Match 47.3%; Score 449.6; DB 2; Length 1679;

Best Local Similarity 73.5%; Pred. No. 5.9e-119;

Matches 613; Conservative 0; Mismatches 194; Indels 27; Gaps 2;

QY 6 CGAGGTCAAGCTGCGAGGATCGAGGGGAGGCTTAGTGCAGCGCTCCCGAAACT 65  
DB 838 CGATATCAAACTGCGAGGCTGCGGCGAAGCTTGAGGTAGGAGCACTCAGTCAAGTT 897  
QY 66 CTCTGTGCGAGGCTCTGATTCATTTCAGTAGCTTTGGAATGACCTGGGCTTCAGTGGC 125  
DB 898 GTCCGCGACAGCTTCTGGCTTCAACATTAAAGATCCTCTATGACACTGGTTAGGACAGG 957  
QY 126 TCCAGAGAAAGGCGTGAAGTGGCTGCATATATTAAGTAGGACAGTATCATCTACTTA 185  
DB 958 GCCTGAACAGGCGCTGAGTGGATTGATGATGATTCCTGAAGATGATGATGATGAATA 1017  
QY 186 TGCACACACAGTAGAGGAGCGATTGACCATCTCCAGAGCAATCCCAAGAAACCCCTGT 245  
DB 1018 TCCCCGAGGCTTCAGGCGAAGGCACTTTTACTACAGACATCTCCAAACAGGCTTA 1077  
QY 246 CCTGCAAAATGACCACTCTAAGGCTGAGGACACAGGTCATGATTAATGCTGCTCTC 293  
DB 1078 CCTGACGCTGACAGCCTGACATCTGAGGACACTGCCCTCTATTATTTGTAATGAGGAGC 1137  
QY 294 -----TGCAGAGATTACGGGGCTTATTGGGGCCAAAGGACCAAGCTCAGCTCTCTC 347  
DB 1138 TCCGACTGGGCGTACTTACTTGTGACTACCTGGGCGCAAGGAGCAAGGCTCCTCTC 1197  
QY 348 AGGTGAGGCGGCTGAGGCGGAGGTGGCTTGGCGGTGGCGGATGGAATGAGCTCAAC 407  
DB 1198 AGGTGAGGCGGCTTGAAGGCGGAGGTGGCTTGGCGGTGGCGGATGGAATGAGCTCAAC 1257  
QY 408 CCAAGTCTCAGCAATCATGTCTGCACTCTCCAGGGGAGAGGCTCAACATGACCTGCAAGTGC 467  
DB 1258 CCAAGTCTCAGCAATCATGTCTGCACTCTCCAGGGGAGAGGCTCAACATGACCTGCAAGTGC 1317  
QY 468 CAGTTCAAGTGAAGTACATGAACCTGTTCCAAACAGAAAGTACAGGACCTCCCAAAAG 527  
DB 1318 CAGCTCAAGTGAAGTACATGAACCTGTTCCAGAGAAAGCAAGGACCTTCCCAAACT 1377  
QY 538 ATGATTTATATGACATCCAACTGCTTCGAGAGTCCCTGCTCCCTTCAAGTGGAGTGG 587  
DB 1378 CTGGATTTATATGACATCCAACTGCTTCGAGAGTCCCTGCTCCCTTCAAGTGGAGTGG 1437  
QY 588 GTCTGGGAGCTCTTACTCTCTCAATCAGAGACAGAGGCTGGAAGATGCTGCCACTTA 647  
DB 1438 ATCTGGGAGCTCTTACTCTCTCAATCAGAGACAGAGGCTGGAAGATGCTGCCACTTA 1497

QY 648 CTACTGCCAGAGTGGAGTGAATCAACCTACCTGTCGTGTGGAGCCAAAGCTGAGCT 707  
DB 1498 TTACTGCCAGCAACGAGTAGTACCCACTCAGCTTCGTTGCTGGACCAAGCTGAGCT 1557  
QY 708 GAAAGGGGCGGCGCGAGAACAAAACTCATCTCAGAAAGATCTGAATGGGCGGTGTA 767  
DB 1558 GAAAGGGGCGGCG-----AGCTCAGCGAGGCGGCGGTAGCGATATCGCGCGCGC 1608  
QY 768 CGAACAAAACTCATCTCAGAAAGATCTGAATGCTGTGGCCAGACAGCA 821  
DB 1609 AGAACAAAACTCATCTCAGAAAGATCTGAATGCGCGCGCGCATCACATCA 1662

RESULT 11

US-09-188-082-15

/ Sequence 15, Application US/09188082

/ Patent No. 6270765

/ GENERAL INFORMATION:

/ APPLICANT: Yashwant M. Deo

/ APPLICANT: Joel Goldstein

/ APPLICANT: Robert Graziano

/ APPLICANT: Chezian Somasundaram

/ TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISD

/ TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES

/ NUMBER OF SEQUENCES: 16

/ CORRESPONDENCE ADDRESS:

/ ADDRESS: LAHIVE & COCKFIELD

/ STREET: 60 State Street, Suite 510

/ CITY: Boston

/ STATE: Massachusetts

/ COUNTRY: USA

/ ZIP: 02109-1875

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: PatentIn Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/188,082

/ FILING DATE:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 08/661,052

/ FILING DATE:

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Arnold, Beth E.

/ REGISTRATION NUMBER: 35,430

/ REFERENCE/DOCKET NUMBER: MXI-043CP

/ TELECOMMUNICATION INFORMATION:

/ TELEFAX: (617)227-5941

/ INFORMATION FOR SEQ ID NO: 15:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 1679 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: cDNA

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 11..1667

US-09-188-082-15

Query Match 47.3%; Score 449.6; DB 3; Length 1679;

Best Local Similarity 73.5%; Pred. No. 5.9e-119;

Matches 613; Conservative 0; Mismatches 194; Indels 27; Gaps 2;

QY 6 CGAGGTCAAGCTGCGAGGATCGAGGGGAGGCTTAGTGCAGCGCTCCCGAAACT 65  
DB 838 CGATATCAAACTGCGAGGCTGCGGCGAAGCTTGAGGTAGGAGCACTCAGTCAAGTT 897  
QY 66 CTCTGTGCGAGGCTCTGATTCATTTCAGTAGCTTTGGAATGACCTGGGCTTCAGTGGC 125

```

Db      898 GTCTGCACAGCTTCTGCTTCAACATTAAGAATCTCTTAATGCACTGCTGAGGACGAG 957
Qy      126 TCACAGAGAGGGGCTGAGTGGGTCGCATATATTAAGTGGCAGATGACATCTACTA 185
Db      958 GCGTGAACAGGGGCTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
Qy      186 TGCAGACACAGTGAAGGAGAGTTCACCATCTTCAGAGACATCCCAAGAGACCCCTGTT 245
Db      1018 TCCCCGAGGTTCCAGGGGCAAGGCCACTTTTACTACAGACATCTCCCAACAGAGCTTA 1077
Qy      246 CCTGCAAAATGACAGCTTAAGCTTGAAGACACAGCTCAATGATTAATG----- 293
Db      1078 CCTGACAGCTGACAGCTTACATCTGAGGACATCGCGCTTATTAATGATGAGGAGAC 1137
Qy      294 -----TGCAGAGATTAAGGGGCTTATTGAGGCGCAAGGAGACCAAGCTGACCTGCTC 347
Db      1138 TCCGAGTGGGCGGCTACTTATTTGATCTGAGGCGCAAGGAGACCAAGCTGACCTGCTC 1197
Qy      348 AGGTGAGGCGGCTCAGGCGGAGTGGCTCTGCGGCTGCGGATCGGACATTTGAGCTCAC 407
Db      1198 AGGTGAGGCGGCTTACGCGGAGTGGCTCTGCGGCTGCGGATCGAAGAAATGCTCTCAC 1257
Qy      408 CCAAGTCTCCAGCAATATGTCTGCAATCTCCAGGAGAGAGGCTCAACATGACCTGACGTGC 467
Db      1258 CCAAGTCTCCAGCAATATGTCTGCAATCTCCAGGAGAGAGGCTCAACATGACCTGACGTGC 1317
Qy      468 CAGTTCAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 527
Db      1318 CAGCTCAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377
Qy      528 ATGATTTATGACATATCCAACTGCTCTGAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 587
Db      1378 CTGGAATTTATGACATATCCAACTGCTCTGAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
Qy      588 GTCTGGAAGCTTCTTACTCTCTCAATCAAGACATGAGAGGCTGAAGATGCTGCCACTTA 647
Db      1438 ATCTGGAAGCTTCTTACTCTCTCAATCAAGACATGAGAGGCTGAAGATGCTGCCACTTA 1497
Qy      648 CTACTGCCAGCAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 707
Db      1498 TTACTGCCAGCAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1557
Qy      708 GAAACGGGCGGCGCAGAAACAAATCTCATCTCAGAAAGAGATCTGAATGAGGCGCTGCA 767
Db      1558 GAAACGGGCGGCGC-----AGGCTGAGCGGAGGCGGCGGAGTGAATGAGGCGGCGC 1608
Qy      768 CGAACAAAATCTCATCTCAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 821
Db      1609 AGAACAAAATCTCATCTCAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1662

```

RESULT 12  
US-09-364-088-15  
Sequence 15, Application US/09364088  
Patent No. 6365161

GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo, et al.  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
OF ANTI-PC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street, 24th Floor  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Releasee #1.0, Version #1.25  
CURRENT APPLICATION DATA:

```

? APPLICATION NUMBER: US/09/364,088
? FILING DATE:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 09/188,082
? FILING DATE: 07-JUNE-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/484,172
? FILING DATE: 07-JUNE-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Remillard, Jane E.
? REGISTRATION NUMBER: 38,972
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617)742-7400
? TELEFAX: (617)742-7414
? INFORMATION FOR SEQ ID NO: 15:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1679 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 11..1667
? US-09-364-088-15

```

```

Query Match 47.3%; Score 449.6; DB 3; Length 1679;
Best Local Similarity 73.5%; Pred. No. 5,9e-119;
Matches 613; Conservative 0; Mismatches 194; Indels 27; Gaps 2;

Qy      6 CAGAGTCAAGTGCAGAGTCAAGGAGGAGCTTATGTCAGCTGAGAGGTCGCCGAAGT 65
Db      838 CGATATCAATGTCAGAGTCTGCGGCAAGACTGTGAGAGTCAAGGACCTGACGTCAAGT 897
Qy      66 CTCTGTGAGGCTTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 125
Db      898 GTCTGCACAGCTTCTGCTTCAACATTAAGAATCTCTTAATGCACTGCTGAGGACGAG 957
Qy      126 TCACAGAGAGGGGCTGAGTGGGTCGCATATTAAGTGGCAGATGACATCTACTA 185
Db      958 GCGTGAACAGGGGCTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
Qy      186 TGCAGACACAGTGAAGGAGAGTTCACCATCTTCAGAGACATCCCAAGAGACCCCTGTT 245
Db      1018 TCCCCGAGGTTCCAGGGGCAAGGCCACTTTTACTACAGACATCTCCCAACAGAGCTTA 1077
Qy      246 CCTGCAAAATGACAGCTTAAGCTTGAAGACACGCTCATGTAATGCTG----- 293
Db      1078 CCTGACAGCTGACAGCTTACATCTGAGGACATCGCGCTTATTAATGATGAGGAGAC 1137
Qy      294 -----TGCAGAGATTAAGGGGCTTATTGAGGCGCAAGGAGACCAAGCTGACCTGCTC 347
Db      1138 TCCGAGTGGGCGGCTACTTATTTGATCTGAGGCGCAAGGAGACCAAGCTGACCTGCTC 1197
Qy      348 AGGTGAGGCGGCTCAGGCGGAGTGGCTCTGCGGCTGCGGATCGGACATTTGAGCTCAC 407
Db      1198 AGGTGAGGCGGCTTACGCGGAGTGGCTCTGCGGCTGCGGATCGAAGAAATGCTCTCAC 1257
Qy      408 CCAAGTCTCCAGCAATATGTCTGCAATCTCCAGGAGAGAGGCTCAACATGACCTGACGTGC 467
Db      1258 CCAAGTCTCCAGCAATATGTCTGCAATCTCCAGGAGAGAGGCTCAACATGACCTGACGTGC 1317
Qy      468 CAGTTCAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 527
Db      1318 CAGCTCAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377
Qy      528 ATGATTTATGACATATCCAACTGCTCTGAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 587
Db      1378 CTGGAATTTATGACATATCCAACTGCTCTGAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
Qy      588 GTCTGGAAGCTTCTTACTCTCTCAATCAAGACATGAGAGGCTGAAGATGCTGCCACTTA 647

```

Db 1438 ATCTGGAGCCTCTTACTCTCTCAATATGAGCGGAATGAGGGCTGAAGATGCTGCCACTTA 1497  
Qy 648 CTACTGCCAGCGTGTGAGTGAATCCCACTCACTGCTGGTGTGGAGCAAGCTGAGCT 707  
Db 1498 TTACTGCGCAGCAACGAGTGTATACCACTCACTGCTGGTGTGGAGCAAGCTGAGCT 1557  
Qy 708 GAAACGGGGCGCGCAGAACAAAATCTCATCTCAGAAAGATCTGAATGGGGCGGTGCA 767  
Db 1558 GAAACGGGGCGCGC-----AGGCTCGAGCGAGGGCGGGGAGCGATATGCGCGCGC 1608  
Qy 768 CGAACAAAACTCATCTCAGAAAGAGATCTGAATGCTGTGGCGCAGACACGCA 821  
Db 1609 AGAACAAAACTCATCTCAGAAAGAGATCTGAATGGCGCGCAGCATCACATCA 1662

RESULT 13  
US-09-102-716-15  
Sequence 15, Application US/09102716  
Patent No. 6395272  
GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo  
Joel Goldstein  
Robert Graziano  
Chezian Somasundaram  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
OF ANTI-FC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/102,716  
FILING DATE: 22-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/661,052  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MXI-043CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1679 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 11..1667  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-102-716-15

Query Match 47.3%; Score 449.6; DB 3; Length 1679;  
Best Local Similarity 73.5%; Pred. No. 5,9e-119;  
Matches 613; Conservative 0; Mismatches 194; Indels 27; Gaps 2;  
Qy 6 CGAGGTCAAGTCTGCAAGAGTCAAGGCGAGGCTTAAGTCAAGCTTGAGAGGCTCCGGAAACT 65  
Db 838 CGATATCAAACTGACAGAGTCTGGGGCAGAACTTGAGAGTCAAGGAGCACTCAAGT 897

Qy 66 CTCCGTGACAGCTCTGGAATTCATTTCAGTAGAGCTTTGGAATGCACTGGGTTGTCAGGC 125  
Db 898 GTCTGCAAGCTTCTGCTTTCATATTAAGATCTCTATATGCACTGGTTGAGCGAGG 957  
Qy 126 TCACAGAAAGGGGCTGAGTGGTGCATATATTAAGTAGAGGAGTGAATCAATCTACTA 185  
Db 958 GCTGTAAAGGGGCTGAGTGGATGGAATTGATGATTTGATCTCTGAGATGAGTGAATCA 1017  
Qy 186 TGCAGACACAGTGAAGGAGCAGATTCAATCTCCAGAGACAAATCCCAAGAACCCCTGT 245  
Db 1018 TGCCCCGAGTTCAGGGCAGAGCCACTTTTACTACAGACATCTCCACACAGACCTTA 1077  
Qy 246 CCTGCAATGACCAAGTCTAAGGTGAGAGACACGGTCAATGATTAAGTCTG----- 293  
Db 1078 CTGCAAGTGAAGAGCTTGACATCTGAGAGACATGCGCTTATTTATTTGATGAGGGAC 1137  
Qy 294 -----TGCAAGATTAACGGGGCTTATTTGGGGCCCAAGGACCAAGTCAAGCTCTC 347  
Db 1138 TCCGACTGGGCGTACTTCTTTGACTACTGGGGCCCAAGGACCAAGTCAAGCTCTC 1197  
Qy 348 AGGTGAGGGCGGCTCAGCGGAGGTGCTGTGGCGGTGGCGGATGGAATTTAGCTCAC 407  
Db 1198 AGGTGAGGGCGGCTCAGCGGAGGTGCTGTGGCGGTGGCGGATGGAATTTAGCTCAC 1257  
Qy 408 CCAAGTCCAGCAATCATGCTGCATCTCCAGGGGAGAGGTCACCATGACCTGAGAGTC 467  
Db 1258 CCAAGTCCAGCAATCATGCTGCATCTCCAGGGGAGAGGTCACCATGACCTGAGAGTC 1317  
Qy 468 CAGTTCAAGTGAAGTATGATGAATGAGTGTCCACAGAGTCAAGGACCTCCCAAGAA 527  
Db 1318 CAGTTCAAGTGAAGTATGATGAATGAGTGTCCACAGAGTCAAGGACCTCCCAAGAA 1377  
Qy 528 ATGATTTATGACATATCAAACTGTCTTCTGAGAGTCCCTGCTTCAAGTGGAGTGG 587  
Db 1378 CTGATTTATGACATATCAAACTGTCTTCTGAGAGTCCCTGCTTCAAGTGGAGTGG 1437  
Qy 588 GTCTGGGACCTCTTACTCTCTCAAAATCAGAGACATGAGGCTGAAGTGTGCGCACTTA 647  
Db 1438 ATCTGGAGCCTCTTACTCTCTCAAAATCAGAGACATGAGGCTGAAGTGTGCGCACTTA 1497  
Qy 648 CTACTGCAGAGTGAAGTGAATCCCACTCACTGCTGTGGTGGAGCAAGCTGAGCT 707  
Db 1498 TTACTGCGCAGCAACGAGTGTATACCACTCACTGCTGTGGTGGAGCAAGCTGAGCT 1557  
Qy 708 GAAACGGGGCGCGCAGAACAAAATCTCATCTCAGAAAGATCTGAATGGGGCGGTGCA 767  
Db 1558 GAAACGGGGCGCGC-----AGGCTCGAGCGAGGGCGGGGTAGCGATTCGCGCGCGC 1608  
Qy 768 CGAACAAAACTCATCTCAGAAAGAGATCTGAATGCTGTGGCGCAGACACGCA 821  
Db 1609 AGAACAAAACTCATCTCAGAAAGAGATCTGAATGGCGCGCAGCATCACATCA 1662

RESULT 14  
US-09-142-974B-3  
Sequence 3, Application US/09142974B  
Patent No. 6451995  
GENERAL INFORMATION:  
APPLICANT: Cheung, Nai-Kong V.  
APPLICANT: Larson, Steven M.  
APPLICANT: Guo, Hong-Fen  
APPLICANT: Rivlin, Ken  
APPLICANT: Sadelain, Michel  
TITLE OF INVENTION: Single chain Fv Constructs of Anti-Ganglioside GD2  
FILE REFERENCE: MSK\_P-013-USNP  
CURRENT APPLICATION NUMBER: US/09/142,974B  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: PCT/US97/04427  
PRIOR FILING DATE: 1997-03-20  
PRIOR APPLICATION NUMBER: 60/013,703  
PRIOR FILING DATE: 1996-03-20  
NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3  
 LENGTH: 1176  
 TYPE: DNA  
 ORGANISM: Murine  
 FEATURE:  
 OTHER INFORMATION: 5F11-scfv-streptavidin  
 NAME/KEY: unsure  
 LOCATION: (37)  
 NAME/KEY: unsure  
 LOCATION: (79)  
 US-09-142-974B-3

Query Match 46.6%; Score 443.6; DB 3; Length 1176;  
 Best Local Similarity 76.6%; Pred. No. 2,6e-117;  
 Matches 558; Conservative 0; Mismatches 161; Indels 9; Gaps 1;

```

QY 8 AGGTCAAGCTGCAGAGTCAGAGGGAGGCTTACTGACGCTGGAGGGTCCGGAACTCT 67
DB 2 AGGTGAACTGCAGAGTCAGAGCTGAACTGGTGNAGCTGGGGCTTCACTGTAAGATAT 61
QY 68 CCTGTGCAGCCTCTGGATTCACTTTCAGTAGAGCTTTGGATGCACTGGGTTCTCAGGCTC 127
DB 62 CCTGCAGAGCTTCTGGAAACAAATTCATGAAATACACCATGCACTGGGTGAAGCAGAGCC 121
QY 128 CAGAGAGGGGCTGAGAGGCTGCAATATATTAATTAAGTGGCAGTAGTACCATCTACTATG 187
DB 122 ATGGAAGAGGCTTGAAGTGGATTTGAGATTTATCTTAACATGGTGGTAACTAACTACA 181
QY 188 CAGACACAGTGAAGGAGCATTCACCATCTCCAGAGACAATCCCAAGAACCCCTGTTCC 247
DB 182 ACAGAGAGTTCAAGGGCAAGCCATTGACTGTAGAACAGTCTCCACAGCAGCTTACA 241
QY 248 TGCAGATGACAGTCTAAGTCTGAGAGACCGCTCATATTAATTAATGAGAGATTAAG 307
DB 242 TGAAGCTCCGAGCTGACATCTGAGAGATTTGCACTTATTAATCTGCAAGAGATACA 301
QY 308 GG-----GCTTATGGGGGCCAAGGAGCACAGGTCACCGTCTCCTCAGGTGAGGCG 358
DB 302 CGGTCCCTTTGCTTACTGGGTCCAGAGGACACCGTCTCTCTCAGGTGAGGCG 361
QY 359 GCTCAGGCGAGGTGCTTGGCGGTGGCGGATCGGACATTAAGTCAACCCAGTCTCCAG 418
DB 362 GTTCAGGCGAGGTGCTTGGCGGTGGCGGATCGGACATTAAGTCAACCCAGTCTCCAG 421
QY 419 CAATCATGCTGATCTCCAGGGGAGAGGCTCACTGACCTGCACTGCAAGTTCAGAGT 478
DB 422 CAATCATGCTGATCTCCAGGGGAGAGGCTCACTGACCTGCACTGCAAGTTCAGAGT 481
QY 479 TAAGATCATGAACTGGTTCAGACAGAGTCAGGACCTCCCAAGAGTGAATTTATG 538
DB 482 TAAGATCATGCACTGGTTCAGACAGAGTCAGGACCTCCCAAGAGTGAATTTATG 541
QY 539 ACACATCAAACTGTCTTCTGAGTCCCTGCTCGCTTCACTGAGTGGAGTGGTCTGGAGCT 598
DB 542 ACACATCAAACTGTCTTCTGAGTCCCTGCTCGCTTCACTGAGTGGAGTGGTCTGGAGCT 601
QY 599 CTTAATCTCTCACAATTCAGAGAGTGAAGTGTGCTGCACTTACTACTGCGACG 658
DB 602 CTTAATCTCTCACAATTCAGAGAGTGAAGTGTGCTGCACTTACTACTGCGACG 661
QY 659 AGTGAAGTGAACCACTCAGGTTGGGCTGGGAGCAAGCTGGAAGTGAAGCGGAGG 718
DB 662 AGGAGAGTGAATCCCGCTCAGCTTGGTGGTGGAGACAGTTGGAATTAACCGGAGG 721
QY 719 CCGCAGAA 726
DB 722 CCGCTGGA 729

```

RESULT 15  
 US-09-473-653-5  
 Sequence 5, Application US/09473653

Patent No. 6703015  
 GENERAL INFORMATION:  
 APPLICANT: Solomon, Beka  
 APPLICANT: Frenkel, Dan  
 TITLE OF INVENTION: IMMUNIZATION AGAINST AMYLOID PLAQUES USING DISPLAY TECHNOLOGY  
 FILE REFERENCE: 00/20785  
 CURRENT APPLICATION NUMBER: US/09/473,653  
 PRIOR FILING DATE: 1999-12-29  
 PRIOR APPLICATION NUMBER: 60/152,417  
 NUMBER OF SEQ ID NOS: 24  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 5  
 LENGTH: 717  
 TYPE: DNA  
 ORGANISM: Artificial sequence  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: (1..1)  
 OTHER INFORMATION: scfv 508f construct  
 US-09-473-653-5

Query Match 46.5%; Score 442; DB 4; Length 717;  
 Best Local Similarity 77.4%; Pred. No. 6e-117;  
 Matches 554; Conservative 0; Mismatches 150; Indels 12; Gaps 1;

```

QY 8 AGGTCAAGCTGCAGAGTCAGAGGGAGGCTTACTGACGCTGGAGGGTCCGGAACTCT 67
DB 2 AGGTGAACTGCAGAGTCAGAGGCTGAGCTGAGTGGGCTTCACTGTAAGATAT 61
QY 68 CCTGTGCAGCCTCTGGATTCACTTTCAGTAGAGCTTTGGATGCACTGGGTTCTCAGGCTC 127
DB 62 CCTGCAGAGCTTCTGGAAACAAATTCATGAAATACACCATGCACTGGGTGAAGCAGAGTC 121
QY 128 CAGAGAGGGGCTGAGAGGCTGCAATATATTAATTAAGTGGCAGTAGTACCATCTACTATG 187
DB 122 ATGGAAGAGGCTTGAAGTGGATTTGAGATTTATCTTAACATGGTGGTAACTAACTACA 181
QY 188 CAGACACAGTGAAGGAGCATTCACCATCTCCAGAGACAATCCCAAGAACCCCTGTTCC 247
DB 182 ACAGAGAGTTCAAGGGCAAGCCATTGACTGTAGAACAGTCTCCACAGCAGCTTACA 241
QY 248 TGCAGATGACAGTCTAAGTCTGAGAGACCGCTCATATTAATTAATGAGAGATTAAG 307
DB 242 TGAAGCTCCGAGCTGACATCTGAGAGATTTGCACTTATTAATCTGCAAGAGATACA 301
QY 308 -----ATTACGGGCTTATTTGGGGCCAAAGGACACCGTCTCTCAGGTGAGG 355
DB 302 CTATGTCTTAATTTGCTACTGGGGCCAAAGTACACCGTCTCTCAGGTGAGG 361
QY 356 GCGGCTCAGGCGAGGTGCTTGGCGGTGGCGGATCGGACATTAAGTCAACCCAGTCTCCAG 415
DB 362 GCGGCTCAGGCGAGGTGCTTGGCGGTGGCGGATCGGACATTAAGTCAACCCAGTCTCCAG 421
QY 416 CAGCAATCATGCTGATCTCCAGGGGAGAGGCTCACTGACCTGCACTGAGTTCAG 475
DB 422 CAGCAATCATGCTGATCTCCAGGGGAGAGGCTCACTGACCTGCACTGAGTTCAG 481
QY 476 GTTAAGATGATGAATCTGGTTCAGACAGAGTCAAGCACTTCCCAAGAGTGAATTT 535
DB 482 GTTAAGATGATGAATCTGGTTCAGACAGAGTCAAGCACTTCCCAAGAGTGAATTT 541
QY 536 ATGACATCAAACTGTCTTCTGAGTCCCTGCTCGCTTCACTGAGTGGAGTGGTCTGGAG 595
DB 542 ATGACATCAAACTGTCTTCTGAGTCCCTGCTCGCTTCACTGAGTGGAGTGGTCTGGAG 601
QY 596 CCTTTACTCTCTCACAATTCAGAGAGTGAAGTGTGCTGCACTTACTACTGCC 655
DB 602 CCTTTACTCTCTCACAATTCAGAGAGTGAAGTGTGCTGCACTTACTACTGCC 661
QY 656 AGCAGTGAAGTGAACCACTCAGGTTGGTGGTGGAGCAAGCTGAGCTGAGAA 711
DB 662 ATCAGCGAGTGAATCCCGCTCAGCTTGGTGGTGGAGCAAGCTGAGAAATTA 717

```

Search completed: June 23, 2005, 10:09:31  
Job time : 204 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 05:50:26 ; Search time 23 Seconds  
(without alignments)  
1321.935 Million cell updates/sec

Title: US-09-403-882a-2

Perfect score: 1622  
Sequence: 1 MAEVKLQESGGGLVPGGSR.....VLTITSLIILMLWQKPR 316

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79;  
1: p1r1;  
2: p1r2;  
3: p1r3;  
4: p1r4;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 892.5 | 55.0        | 268    | 2 A56446 | Ig heavy chain V r |
| 2          | 718.5 | 44.3        | 249    | 2 S41374 | single chain Fv an |
| 3          | 626.5 | 38.6        | 233    | 2 UCS322 | p53 specific singl |
| 4          | 555.5 | 34.2        | 213    | 2 S68213 | Ig heavy chain (Ma |
| 5          | 535   | 33.0        | 136    | 1 G1M621 | Ig heavy chain pre |
| 6          | 522.5 | 32.2        | 121    | 2 B34871 | Ig heavy chain V r |
| 7          | 516   | 31.8        | 130    | 1 JL0079 | Ig kappa chain pre |
| 8          | 514   | 31.7        | 235    | 2 S25058 | Ig kappa chain - m |
| 9          | 511   | 31.5        | 107    | 2 A30562 | Ig kappa chain V r |
| 10         | 508   | 31.3        | 103    | 2 S29591 | Ig kappa chain V r |
| 11         | 508   | 31.3        | 104    | 2 B49049 | Ig kappa chain V r |
| 12         | 506   | 31.2        | 106    | 2 PS0071 | Ig kappa chain V r |
| 13         | 506   | 31.2        | 107    | 2 B30562 | Ig kappa chain V r |
| 14         | 501   | 30.9        | 107    | 2 S11118 | Ig kappa chain V r |
| 15         | 498   | 30.7        | 107    | 2 S11119 | Ig kappa chain V r |
| 16         | 498   | 30.7        | 107    | 2 PC4405 | Ig kappa chain V r |
| 17         | 490   | 30.2        | 107    | 2 S11121 | Ig kappa chain V r |
| 18         | 488   | 30.1        | 107    | 2 PT0406 | Ig kappa chain V r |
| 19         | 484.5 | 29.9        | 143    | 2 S23624 | Ig heavy chain V r |
| 20         | 483   | 29.8        | 107    | 2 S11117 | Ig kappa chain V r |
| 21         | 482   | 29.7        | 120    | 2 S12853 | Ig kappa chain V r |
| 22         | 481.5 | 29.7        | 108    | 2 G30560 | Ig heavy chain V r |
| 23         | 480   | 29.6        | 94     | 2 D25931 | Ig heavy chain V r |
| 24         | 479   | 29.5        | 100    | 2 S29590 | Ig kappa chain V r |
| 25         | 478   | 29.5        | 107    | 2 S11112 | Ig kappa chain V r |
| 26         | 476   | 29.3        | 108    | 2 PH1015 | Ig heavy chain V r |
| 27         | 473   | 29.2        | 107    | 2 PD0011 | Ig kappa chain V r |
| 28         | 473   | 29.2        | 97     | 2 PH1084 | Ig light chain V r |
| 29         | 473   | 29.2        | 114    | 2 S46392 | Ig heavy chain V r |

|    |       |      |     |          |                    |
|----|-------|------|-----|----------|--------------------|
| 30 | 471.5 | 29.1 | 123 | 2 S26794 | Ig heavy chain V r |
| 31 | 471.5 | 29.1 | 125 | 2 S30531 | Ig heavy chain V r |
| 32 | 471   | 29.0 | 114 | 2 S46391 | Ig heavy chain V r |
| 33 | 470   | 29.0 | 110 | 2 PH1014 | Ig heavy chain V r |
| 34 | 470   | 29.0 | 114 | 2 S46390 | Ig heavy chain V r |
| 35 | 469   | 28.9 | 107 | 2 S11113 | Ig kappa chain V r |
| 36 | 468.5 | 28.9 | 548 | 2 S3864  | Ig epsilon chain C |
| 37 | 468   | 28.9 | 97  | 2 S26341 | Ig light chain V r |
| 38 | 467   | 28.8 | 106 | 2 G27887 | Ig kappa chain V r |
| 39 | 466   | 28.7 | 140 | 2 S70442 | Ig heavy chain pre |
| 40 | 463   | 28.5 | 106 | 2 B54378 | Ig light chain V r |
| 41 | 462   | 28.5 | 104 | 2 UC6076 | anti-D-dimer monoc |
| 42 | 461.5 | 28.5 | 108 | 2 S38720 | Ig light chain V r |
| 43 | 461.5 | 28.5 | 120 | 2 S55536 | Ig heavy chain V r |
| 44 | 461   | 28.4 | 130 | 2 S04573 | Ig kappa chain pre |
| 45 | 460.5 | 28.4 | 120 | 2 S66536 | Ig light chain V r |

#### ALIGNMENTS

##### RESULT 1

A56446 Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)

C/Species: Mus musculus (house mouse)

C/Date: 19-Jan-1996 #sequence\_rev1510n 19-Jan-1996 #text\_change 16-Aug-1996

C/Accession: A56446

R/Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.

J. Biol. Chem. 270, 7829-7835, 1995

A/Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical

A/Reference number: A56446; PMID:95229583; PMID:7713873

A/Accession: A56446

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-268 <TRAN>

A/Cross-references: GB:U20617

C/Keywords: heterotetramer; immunoglobulin

Query Match 55.0%; Score 892.5; DB 2; Length 268;  
Best Local Similarity 65.9%; Pred. No. 1.2e-52;  
Matches 176; Conservative 26; Mismatches 52; Indels 13; Gaps 2;

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MAEVKLQESGGGLVPGGSRKLSCAASGFTSSFGHMYRQAEKGLVAVYISGSSSTI    | 60  |
| DB | 1   | MAQVKLQESGAEIVKPGASVKLSCTTSGEFIKDYHVMWVQREQGLEWIGRIAPANGIT | 60  |
| QY | 61  | YVADTVKGRFTISRDNPXKTLFLQMTSLASSEDVMTYCA---RDYGAVYGQGTYYVSS | 116 |
| DB | 61  | KDPRKQKATTAADYSNTRYVQLSLTSEDPVAVYCAVYLLRYENWVGQITVTVSS     | 120 |
| QY | 117 | GGGGSGGGSGGGSGSDIELTQSPAIMSASPERVVTTCASSSVRYVMNMFQKSGTSPKR | 176 |
| DB | 121 | GGGGSGGGSGGGSGSDIELTQSPAIMSASPERVVTTCASSSVRYVMNMFQKSGTSPKR | 180 |
| QY | 177 | WIYDTKLSGGVPAFPSSGSGTYSGLTISMEADPAATYYCOQWSNPLTFGAGTKLEL   | 236 |
| DB | 181 | WVYVTSHPGVPARFSSGSGNSYSGLTISMEGDAATYYCOQFSSPTFSGTGLLEI     | 240 |
| QY | 237 | KRA-----AAEQKLTISEDLNGA                                    | 254 |
| DB | 241 | KRSAAHHHHRGAABQKLTISEDLNGA                                 | 267 |

##### RESULT 2

S41374 single chain Fv antibody - mouse

C/Species: Mus musculus (house mouse)

C/Date: 06-Jan-1995 #sequence\_rev1510n 06-Jan-1995 #text\_change 06-Jan-1995

C/Accession: S41374

R/Artsenko, O.; Weiler, E.W.; Muenz, K.; Conrad, U.

submitted to the EMBL Data Library, January 1994  
A/Description: Construction and functional characterization of a single chain Fv antibody  
A/Reference number: S41374



A:Accession: S41374  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-249 <ART>  
A:Cross-references: EMBL:Z29480

Query Match 44.3%; Score 718.5; DB 2; Length 249;  
Best Local Similarity 57.8%; Pred. No. 4.7e-41;  
Matches 144; Conservative 31; Mismatches 63; Indels 11; Gaps 3;

QY 3 EYKLOESGGGLVOPGSGSRKLSCAASGFTSSFCGMHWVROAPEKGLEWVAYISSGSSITYY 62  
DB 1 QVQLQSGGALVPFGSGSVKLSCTASGPNKDYIHWKQPEKGLMARIAPASGNVXY 60  
QY 63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVWYCARD---YGA--YWGQGTITVVS 117  
DB 61 VPRFQOKATITADTSNTAVYLLSLTSEDTAVYCARDTLVTSIGWQGSSTVVS 120  
QY 118 GGGSGGGGSDIELTQSPAIMSASPERVMTCSASSV-----RYMWFQOKSG 171  
DB 121 GGGSGGGGSDIELTQSPSVVVIPEGSVISCRSSKSLYSDDGYLFWFLQRP 180  
QY 172 TSPKRWIYDTSKSGVPAFSGSGTSYSLTISMEADATYYCOQMSNPLTFGAG 231  
DB 181 QSFQILLIKMSNLASGVPRFSGSGSTFTLRISRVEADGVYICMQRREYPLTFGAG 240  
QY 232 TKLELRRAA 240  
DB 241 TKLELRRAA 249

RESULT 3  
JC5322  
p53 specific single-chain antibody Pab421 - human

C:Species: Homo sapiens (man)  
C:Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
C:Accession: JC5322  
R:Jannot, C.B.; Hynes, N.E.  
Biochem. Biophys. Res. Commun. 230, 242-246, 1997  
A:Title: Characterization of scFv-421, a single-chain antibody targeted to p53.  
A:Reference number: JC5322; MUID:97168950; PMID:9016757  
A:Accession: JC5322  
A:Molecule type: mRNA  
A:Residues: 1-233 <UAN>  
A:Experimental source: hybridoma cell  
C:Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 38.6%; Score 626.5; DB 2; Length 233;  
Best Local Similarity 55.3%; Pred. No. 6.1e-35;  
Matches 131; Conservative 29; Mismatches 66; Indels 11; Gaps 5;

QY 7 QESGGGLVOPGSGSRKLSCAASGFTSSFCGMHWVROAPEKGLEWVAYI--SSGSSITYYAD 64  
DB 1 QESGALVNSGASVYKLSCTTSGFNINDYMWYKRRPEGLWIGRIDENGDADMTRES 60  
QY 65 TVKGRFTISRDNPKNTLFLQMTSLRSEDTVWYCARDYGAWQGTITVVS 124  
DB 61 GYVA--TMTADTSSNTAVYLLSLTSEDTAVYCNAGMD--YWGQGTITVVS 117  
QY 125 GGGGGSDIELTQSPAIMSASPERVMTCSASSV-----RYMWFQOKSGTSPKRWIY 179  
DB 118 ASGGGSDIELTQSPSLAVSLGQRTATISCRASKSVSTSGYSVMHNNQKPKGPPRLLY 177  
QY 180 DTSKLSGGVPAFSGSGTSYSLTISMEADATYYCOQMSNPLTFGAGTKLEI 236  
DB 178 LVSNLESQVPAFSGSGSTFTLRISRVEADGVYICMQRREYPLTFGAG 233

RESULT 4  
S68213  
Ig heavy chain (Mab03-1) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 29-Jul-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004

C:Accession: S68213  
R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.  
FEBS Lett. 375, 273-276, 1995  
A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin I  
A:Reference number: S68211; MUID:96085223; PMID:7498516  
A:Accession: S68213  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-213 <TAK>  
A:Cross-references: UNIPROT:Q91Z05; EMBL:D29667  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:137-201/Domain: immunoglobulin homology <IMM>

Query Match 34.2%; Score 555.5; DB 2; Length 213;  
Best Local Similarity 53.9%; Pred. No. 3.1e-30;  
Matches 124; Conservative 17; Mismatches 50; Indels 39; Gaps 5;

QY 4 VTLQESGGGLVOPGSGSRKLSCAASGFTSSFCGMHWVROAPEKGLEWVAYISSGSSITYYA 63  
DB 2 VQLVESGGGLVOPGSGSRKLSCAASGFTSSFCGMHWVROAPEKGLEWVAYISSGSSITYYA 61  
QY 64 DTVKGRFTISRDNPKNTLFLQMTSLRSEDTVWYCARDY---GAYWQGTITVVS 120  
DB 62 DTVKGRFTISRDNPKNTLFLQMTSLRSEDTAVYCARSLPFDYWGQGTITVVS 118  
QY 121 SGGSGGGGSDIELTQSPAIMSASPERVMTCSASSVRYMWFQOKSGTSPKRWIYD 180  
DB 119 -----KTPPSSVYPLAPGCGDTTGSVTLGLGVKGFPESEVTW-----WN 159  
QY 181 TSKLSSGV---PARFSGSGTSYSLTISMEADATYYCOQMSNPLT 227  
DB 160 SGLSSVHTFPALLQSGLYTWSSSVTVPS-----STWPSQTVT 198

RESULT 5  
GIMS21

Ig heavy chain precursor V region (MOPC 21) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 09-Jul-2004  
C:Accession: E90809; A93184; A02066  
R:Botwell, A.L.M.; Peakel, M.; Reih, M.; Imanishi-Kari, T.; Rajewsky, K.; Baltimore, D.  
Cell 24, 625-637, 1981  
A:Title: Heavy chain variable region contribution to the NP(b) family of antibodies: some  
A:Reference number: A90809; MUID:81234548; PMID:6788376  
A:Accession: E90809  
A:Molecule type: mRNA  
A:Residues: 1-136 <BOT>

A:Cross-references: UNIPROT:P01783; GB:J00522; NID:g195052; PIDN:AAD15290.1; PID:g195055  
R:Adelcupo, K.; Milstein, C.; Secher, D.S.  
Nature 265, 299-304, 1977  
A:Title: Molecular analysis of spontaneous somatic mutants.  
A:Reference number: A93184; MUID:77100368; PMID:401950  
A:Contents: myeloma protein MOPC 21  
A:Accession: A93184

A:Molecule type: protein  
A:Residues: 17-74, 'D', 76-77, 'H', 79-88, 'ND', 91-114, 'H', 116-119, 'W', 121-136 <ADE>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterocyclamer; immunoglobulin  
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>  
F:17-136/Domain: Ig heavy chain V region (MOPC 21) #status experimental <MAT>  
F:115-119/Region: D segment  
F:120-136/Region: D segment (UH4)  
F:38-112/Disulfide bonds: #status experimental

Query Match 33.0%; Score 535; DB 1; Length 136;  
Best Local Similarity 87.5%; Pred. No. 4.4e-29;  
Matches 105; Conservative 5; Mismatches 4; Indels 6; Gaps 2;

QY 3 EYKLOESGGGLVOPGSGSRKLSCAASGFTSSFCGMHWVROAPEKGLEWVAYISSGSSITYY 62  
DB 17 DVQLVESGGGLVOPGSGSRKLSCAASGFTSSFCGMHWVROAPEKGLEWVAYISSGSSITLHY 76





C:/Date: 19-Mar-1997 #sequence\_revision 18-Jul-1997 #text\_change 21-Jan-2000  
 C:/Accession: S1118  
 R:/Kaartinen, M.; Grifflths, G.M.; Markham, A.F.; Milestein, C.  
 Nature 304, 320-324, 1983  
 A:/Title: mRNA sequences define an unusually restricted IgG response to 2-phenylloxazalone  
 A:/Reference number: S07331; M0ID:83271467; PMID:6877353  
 A:/Accession: S1118  
 A:/Status: preliminary  
 A:/Molecule type: mRNA  
 A:/Residues: 1-107 <KAA>  
 C:/Superfamily: Immunoglobulin V region; immunoglobulin homology  
 F:/6-89/Domain: Immunoglobulin homology <IMM>

Query Match 30.7%; Score 501; DB 2; Length 107;  
 Best Local Similarity 91.5%; Pred. No. 6,4e-27;  
 Matches 97; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 133 IELTQSPALMSASPGERVTMTCSASSSVRYMNFQOKSGTSPKRWIYDTSKLSGVPARF 192  
 |||||  
 DB 2 IYLTQSPALMSASPGERVTMTCSASSSVRYMNFQOKSGTSPKRWIYDTSKLSGVPARF 61  
 |||||

QY 193 SGGSGGTSLTSSMEAPDAITYCCQWSSNPLTFGAGTKLELR 238  
 |||||  
 DB 62 SGGSGGTSLTSSMEAPDAITYCCQWSSNPLTFGAGTKLELR 107  
 |||||

## RESULT 15

S1119  
 I:/Kappa chain V region (clone NO6-8.3.1) - mouse (fragment)  
 C:/Species: Mus musculus (house mouse)  
 C:/Date: 19-Mar-1997 #sequence\_revision 23-Aug-1997 #text\_change 21-Jan-2000  
 C:/Accession: S1119  
 R:/Kaartinen, M.; Grifflths, G.M.; Markham, A.F.; Milestein, C.  
 Nature 304, 320-324, 1983  
 A:/Title: mRNA sequences define an unusually restricted IgG response to 2-phenylloxazalone  
 A:/Reference number: S07331; M0ID:83271467; PMID:6877353  
 A:/Accession: S1119  
 A:/Molecule type: mRNA  
 A:/Residues: 1-107 <NAT>  
 C:/Superfamily: Immunoglobulin V region; immunoglobulin homology  
 F:/6-89/Domain: Immunoglobulin homology <IMM>

Query Match 30.7%; Score 498; DB 2; Length 107;  
 Best Local Similarity 89.6%; Pred. No. 1e-26;  
 Matches 95; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 133 IELTQSPALMSASPGERVTMTCSASSSVRYMNFQOKSGTSPKRWIYDTSKLSGVPARF 192  
 |||||  
 DB 2 IYLTQSPALMSASPGERVTMTCSASSSVRYMNFQOKSGTSPKRWIYDTSKLSGVPARF 61  
 |||||

QY 193 SGGSGGTSLTSSMEAPDAITYCCQWSSNPLTFGAGTKLELR 238  
 |||||  
 DB 62 SGGSGGTSLTSSMEAPDAITYCCQWSSNPLTFGAGTKLELR 107  
 |||||

Search completed: June 23, 2005, 05:50:58  
 Job time : 24 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 05:50:26 ; Search time 83 seconds  
(without alignments)  
1949.603 Million cell updates/sec

Title: US-09-403-882a-2  
Perfect score: 1622  
Sequence: 1 MAEVKLGESGGGLVQPGSR.....VLLTISLILMLMOKKPR 316

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID        | Description            |
|------------|-------|-------------|--------|--------------|------------------------|
| 1          | 893   | 55.1        | 240    | 2 Q65ZC9     | Q65ZC9 homo sapien     |
| 2          | 853.5 | 52.6        | 298    | 2 Q9QYF0     | Q9QYF0 synthetic c     |
| 3          | 850   | 52.4        | 255    | 2 Q6K805     | Q6K805 mus musculu     |
| 4          | 790.5 | 48.7        | 248    | 2 Q65ZD7     | Q65ZD7 mus sp. b3(     |
| 5          | 733   | 45.2        | 244    | 2 Q65ZC8     | Q65ZC8 homo sapien     |
| 6          | 707   | 43.6        | 241    | 2 Q65ZL2     | Q65ZL2 mus sp. fv/     |
| 7          | 695   | 42.8        | 241    | 2 Q921A6     | Q921A6 mus musculu     |
| 8          | 658   | 40.6        | 243    | 2 Q7TQW2     | Q7TQW2 mus musculu     |
| 9          | 608.5 | 37.5        | 218    | 2 Q925S1     | Q925S1 mus musculu     |
| 10         | 543   | 33.5        | 485    | 2 Q6PDB8     | Q6PDB8 mus musculu     |
| 11         | 542   | 33.4        | 473    | 2 Q91Z05     | Q91Z05 mus musculu     |
| 12         | 535   | 33.0        | 136    | 1 HV16_MOUSE | HV16_MOUSE mus musculu |
| 13         | 507   | 31.3        | 107    | 1 KV6F_MOUSE | KV6F_MOUSE mus musculu |
| 14         | 506   | 31.2        | 107    | 1 KV6I_MOUSE | KV6I_MOUSE mus musculu |
| 15         | 502   | 30.9        | 107    | 1 KV6G_MOUSE | KV6G_MOUSE mus musculu |
| 16         | 500   | 30.8        | 107    | 2 KV6I_MOUSE | KV6I_MOUSE mus musculu |
| 17         | 498   | 30.7        | 112    | 2 Q6K1F0     | Q6K1F0 mus musculu     |
| 18         | 494   | 30.5        | 107    | 1 KV6J_MOUSE | KV6J_MOUSE mus musculu |
| 19         | 484.5 | 29.9        | 487    | 2 Q99K44     | Q99K44 mus musculu     |
| 20         | 483   | 29.8        | 134    | 2 Q8VDD0     | Q8VDD0 mus musculu     |
| 21         | 480   | 29.6        | 112    | 2 Q8K1F2     | Q8K1F2 mus musculu     |
| 22         | 473.5 | 29.2        | 464    | 2 Q6MZ66     | Q6MZ66 homo sapien     |
| 23         | 473   | 29.2        | 472    | 2 Q6N0B9     | Q6N0B9 homo sapien     |
| 24         | 468.5 | 28.9        | 473    | 2 Q6MZV7     | Q6MZV7 homo sapien     |
| 25         | 462.5 | 28.5        | 113    | 2 Q9UL90     | Q9UL90 homo sapien     |
| 26         | 462   | 28.5        | 478    | 2 Q6P1R1     | Q6P1R1 homo sapien     |
| 27         | 462   | 28.4        | 613    | 2 Q8MUK1     | Q8MUK1 homo sapien     |
| 28         | 461   | 28.3        | 112    | 2 Q8K1F3     | Q8K1F3 mus musculu     |
| 29         | 459   | 28.3        | 470    | 2 Q6PJ44     | Q6PJ44 homo sapien     |
| 30         | 458   | 28.2        | 114    | 2 Q8K1F1     | Q8K1F1 mus musculu     |
| 31         | 454.5 | 28.0        | 597    | 2 Q96BB9     | Q96BB9 homo sapien     |

|    |       |      |     |              |                        |
|----|-------|------|-----|--------------|------------------------|
| 32 | 452.5 | 27.9 | 119 | 2 Q920E7     | Q920E7 mus musculu     |
| 33 | 450   | 27.7 | 606 | 2 Q6GMV2     | Q6GMV2 homo sapien     |
| 34 | 449   | 27.7 | 466 | 2 Q6IN78     | Q6IN78 homo sapien     |
| 35 | 446   | 27.5 | 107 | 1 KV6C_MOUSE | KV6C_MOUSE mus musculu |
| 36 | 445   | 27.4 | 118 | 2 Q9UL51     | Q9UL51 homo sapien     |
| 37 | 443   | 27.3 | 465 | 2 Q6P6C4     | Q6P6C4 homo sapien     |
| 38 | 443   | 27.3 | 479 | 2 Q91WP5     | Q91WP5 mus musculu     |
| 39 | 441   | 27.2 | 107 | 1 KV6B_MOUSE | KV6B_MOUSE mus musculu |
| 40 | 440.5 | 27.2 | 116 | 2 Q9UL53     | Q9UL53 homo sapien     |
| 41 | 440   | 27.1 | 107 | 1 KV6D_MOUSE | KV6D_MOUSE mus musculu |
| 42 | 439.5 | 27.1 | 484 | 2 Q96K68     | Q96K68 homo sapien     |
| 43 | 439   | 27.0 | 466 | 2 Q91Z07     | Q91Z07 mus musculu     |
| 44 | 438.5 | 27.0 | 131 | 2 Q811C3     | Q811C3 mus musculu     |
| 45 | 438.5 | 27.0 | 544 | 2 Q6PJ95     | Q6PJ95 homo sapien     |

## ALIGNMENTS

| RESULT 1   | Q65ZC9   | PRELIMINARY; | PRT; | 240 AA. |
|--|--|--------------|------|---------|
| ID   | Q65ZC9   |              |      |         |
| AC   | Q65ZC9   |              |      |         |
| DT   | 25-OCT-2004 (TREMBLrel. 28, Created)                               |              |      |         |
| DT   | 25-OCT-2004 (TREMBLrel. 28, last sequence update)                  |              |      |         |
| DE   | 25-OCT-2004 (TREMBLrel. 28, last annotation update)                |              |      |         |
| DE   | Single-chain Fv (Fragment).  |              |      |         |
| GN   | Name=ScFv;   |              |      |         |
| OS   | Homo sapiens (Human).  |              |      |         |
| OC   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |              |      |         |
| OC   | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.        |              |      |         |
| OX   | NCBI_TaxID=9606;   |              |      |         |
| RN   | [1]  |              |      |         |
| RP   | SEQUENCE FROM N.A.   |              |      |         |
| RC   | STRATIN=CIG/7;   |              |      |         |
| RA   | MEDLINE=97362799; PubMed=9219263;                                  |              |      |         |
| RX   | Kontermann R.E., Wing M.G., Winter G.;                             |              |      |         |
| RT   | "Complement recruitment using bispecific diabodies";               |              |      |         |
| RL   | Nat. Biotechnol. 15:629-631(1997).                                 |              |      |         |
| DR   | EMBL: Y13056; CAA73499.1; -  |              |      |         |
| DR   | InterPro: IPR003599; IG_   |              |      |         |
| DR   | InterPro: IPR007110; IG_1ike.                                      |              |      |         |
| DR   | InterPro: IPR003596; IG_v.   |              |      |         |
| DR   | Pfam: PF00047; IG_2.   |              |      |         |
| DR   | SMART: SM00409; IG_2.  |              |      |         |
| DR   | SMART: SM00406; IG_2.  |              |      |         |
| DR   | PROSITE: PS50835; IG_1ike; 2.                                      |              |      |         |
| FT   | NON_TER 1  |              |      |         |
| FT   | NON_TER 240  |              |      |         |
| SQ   | SEQUENCE 240 AA; 25569 MW; FDCFD3645F64B373 CRC64;                 |              |      |         |
| Query Match  | 55.1%; Score 893; DB 2; Length 240;                                |              |      |         |
| Best Local Similarity  | 70.4%; Pred. No. 1.3e-57;  |              |      |         |
| Matches 169; Conservative 32; Mismatches 35; Indels 4; Gaps 2; |  |              |      |         |
| QY   | 3 EVKLQESGGGLVQPGSGRLKSLCAAGFTFSSFGMHVWROAPEKLEWVAIYSSGSSITYY 62   |              |      |         |
| DB   | 1 QVQLVQSGGGLVPGGSLRLKSLCAAGFTFSSYGMHWROAPKGLEWVAIYSDGSKITY 60     |              |      |         |
| QY   | 63 ADTVKRGFTISRDNPKNTFLQNTSLRSEDVTWYYCARDYGV--WGQTTVTYSSGGG 119    |              |      |         |
| DB   | 61 ADSVKGFTISRDNPKNTFLQNTSLRSEDVTWYYCARDYGV--WGQTTVTYSSGGG 120     |              |      |         |
| QY   | 120 GSGGGSGGGGSLIELTQSPALMSASGERVTMTCASSSV--TYMMFQOKSGTSPKRWI 178  |              |      |         |
| DB   | 121 GSGGGSGGGGSLIQMTQSPSTLSASISGRVTTCRASISGIVYMWLAAYQKPGAPFLTI 180 |              |      |         |
| QY   | 179 YDPSKLSGVPARFSGSGSTYSLTISWEARDAATYYCOQMSNPLTGAQTKLEIKR 238     |              |      |         |
| DB   | 181 YKASLSLASRAPRFSGSGSTDTLTITISLPDDFATYYCOQISNPLTGGGKLEIKR 240    |              |      |         |

| ID | Q9QYFO   | PRELIMINARY;                            | PRT; | 298 AA. |
|----|--|---|------|---------|
| AC | Q9QYFO   |   |      |         |
| DT | 01-MAY-2000  | (TREMBLrel. 13, Created)                |      |         |
| DT | 01-MAY-2000  | (TREMBLrel. 13, Last sequence update)   |      |         |
| DE | 01-OCT-2003  | (TREMBLrel. 25, Last annotation update) |      |         |
| DE | CN 8 single chain antibody.                                      |   |      |         |
| GN | Name=CN 8 scFv;  |   |      |         |
| OS | synthetic construct.   |   |      |         |
| OC | other sequences; artificial sequences.                           |   |      |         |
| OX | NCBI_Taxid=32630;  |   |      |         |
| RN | [1]  |   |      |         |
| RP | SEQUENCE FROM N.A.   |   |      |         |
| RX | MEDLINE=20183931; PubMed=10706631; DOI=10.1073/pnas.050582197;   |   |      |         |
| RA | Shinohara N., Demura T., Fukuda H.;                              |   |      |         |
| RT | "Isolation of a vascular cell wall-specific monoclonal antibody  |   |      |         |
| RT | recognizing a cell polarity by using a phase display subtraction |   |      |         |
| RT | method."   |   |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).                |   |      |         |
| DR | EMBL; AB036341; BAA86833.1; -.                                   |   |      |         |
| DR | PIR; A33933; A33933.   |   |      |         |
| DR | PIR; S19112; S19112.   |   |      |         |
| DR | HSSP; P01820; 1A70.  |   |      |         |
| DR | InterPro; IPR007110; Ig-like.                                    |   |      |         |
| DR | InterPro; IPR003596; Ig_v.                                       |   |      |         |
| DR | SMART; SM00406; IGV; 2.  |   |      |         |
| DR | PROSITE; PSS0835; IG_LIKE; 2.                                    |   |      |         |
| QO | SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;               |   |      |         |

| Query Match | Best Local Similarity | 52.6%   | Score 853.5         | DB 2   | Length 298 |
|-------------|-----------------------|---|---------------------|--------|------------|
| Matches 163 | Conservative 32       | Mismatches 46   | Indels 5            | Gaps 2 |            |
| Qy          | 1                     | MAEYKLOEGSGGGLVOPGGSRKLSCAAGFTBSSFGMHVROAPFEGLEWVAYISSGSSIT   | 60                  |        |            |
| Db          | 38                    | MAQYKLOQSGGGLVXPGGSLKSCAAGSDPFRYMSWVROAPGKLEWIGINPDSSTI       | 97                  |        |            |
| Qy          | 61                    | YYADTVKGRFTISRDNPNKNTLFLQWTSLSREPTVWYYCAR----                 | DYGAWVGQGTIVYSS 116 |        |            |
| Db          | 98                    | NYTSLNDKFLIISRDNAKNTLYLQMSKVRSEDTALYYCARASYGHSAVGGGTTIVYSS    | 157                 |        |            |
| Qy          | 117                   | GGGGSGGGSGGGSGGSDIEFLTOSPALMSAPGERVMTCSASSV-RYNNFPOKSGTSPK    | 175                 |        |            |
| Db          | 158                   | GGGGSGGGSGGGSGGSDIELTOSPALSASVGETVITTORASGINHNYLAWYQOKGKSPQ   | 217                 |        |            |
| Qy          | 176                   | RWIDTSKLSGGVPAFRFSGSGSGTYSYSLTISMEAEADATYATCCQOWSNPLTFGAGTKLE | 235                 |        |            |
| Db          | 218                   | LIVYNAKTLACGVSRFSGSGSGGYSLKINSLOGEDGSGYCCQHFMTPTPYFGGKYLE     | 277                 |        |            |
| Qy          | 236                   | LKRAAA 241  |                     |        |            |
| Db          | 278                   | IKRAAA 283  |                     |        |            |

RESULT 3

Q6K605 PRELIMINARY; PRT; 255 AA.

AC Q6K605;

DT 05-JUL-2004 (TREMBLrel. 27, Created)

DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE SCFV B8E5 protein (Fragment).

GN CN

GN Name=SCFV B8E5;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RA [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Balb/c;

RA Peter J.C., Wallukat G., Tugler J., Maurice D., Roegel J.C.,

RA Brian J.P., Hoebeke J.; the EMBL/GenBank/DBJ databases.

RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ746180; CAG34081.1; -.  
DR HSSP; P01837; 1KCR.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003596; Ig\_v.  
DR SMART; SM00409; IG; 2.  
DR SMART; SM00406; IGv; 2.  
DR PROSITE; PSS0835; IG\_LIKE; 2.  
DR NON TER 1  
SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

| Query Match           | 52.4%            | Score 850;  | DB 2;      | Length 255; |
|-----------------------|------------------|---|------------|-------------|
| Best Local Similarity | 67.1%            | Pred. No. 2e-54;  |            |             |
| Matches 167;          | Conservative 29; | Mismatches 39;  | Indels 14; | Gaps 4;     |
| QY                    | 3                | EYVLQDSGGGLVQPGSGRKLSCAASGTFSSFGCMHWYRQAPDKCLEWVAIYISGSSSTIYY | 62         |             |
| Db                    | 1                | QVLOQDSGGGLVHPFGSLKVCASASGTFSSYSGMSWRQTPDKCLEWVAITISGGSYTY    | 60         |             |
| QY                    | 63               | ADTVKGRFTISRNPKNKTLFLQMTSLASEDTWMYCAR--DY---GA--YWGQGTIVVS    | 115        |             |
| Db                    | 61               | PDSVKGKFTISRNNAKNTLYLQMSLSKSEDTWAMYCARHINRYDAFDYWGQGTITLVS    | 120        |             |
| QY                    | 116              | SGGGSGGGGSGGGGSPDIELTQSPALMSAPRGERTMTGASSV-----RMMNFQ         | 168        |             |
| Db                    | 121              | SGGGSGGGGSGGGGSDIVMAQSPSSLSVAGEVIMSSCSSOSLMSRNOKNLWAYQ        | 180        |             |
| QY                    | 169              | KSGTSPKRWIYDTSKLSGVPARFSGSGSTSYSLTISMEABDPAATYCCQWSSNP        | 228        |             |
| Db                    | 181              | KPGQSPKLILYIGASTRESGVDPDRFTGSGSTDFTLTISVQAEADLAVYYCQNDHSYELTF | 240        |             |
| QY                    | 229              | GAGTKLEIK   | 237        |             |
| Db                    | 241              | GAGTKLEIK   | 249        |             |

| RESULT 4 |  |              |              |
|----------|--|--------------|--------------|
| ID       | 065Z07   | PRELIMINARY; | PRT; 248 AA. |
| AC       | 065Z07;  |              |              |
| DT       | 25-OCT-2004 (TREMBLrel. 28, Created)                               |              |              |
| DT       | 25-OCT-2004 (TREMBLrel. 28, Last sequence update)                  |              |              |
| DT       | 25-OCT-2004 (TREMBLrel. 28, Last annotation update)                |              |              |
| DE       | B3 (Fv)-PE40 (Fragment).   |              |              |
| GN       | Name=B3 (Fv) -PE40;  |              |              |
| OS       | Mus sp.  |              |              |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |              |              |
| OC       | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |              |              |
| OX       | NCBI_TaxID=10095;  |              |              |
| RN       | [1]  |              |              |
| RP       | SEQUENCE FROM N.A.   |              |              |
| RX       | MEDLINE=92020904; PubMed=1924323;                                  |              |              |
| RA       | Brinkmann U., Pai L.H., Fitzgerald D.J., Willingham M., Pastan I., |              |              |
| RT       | "B3 (Fv)-PE38KDEL, a single-chain immunotoxin that causes complete |              |              |
| RT       | regression of a human carcinoma in mice."                          |              |              |
| RL       | Proc. Natl. Acad. Sci. U.S.A. 88:8616-8620(1991).                  |              |              |
| DR       | EMBL; S57990; AAB1997.2; -.  |              |              |
| DR       | InterPro; IPR003599; IG.   |              |              |
| DR       | InterPro; IPR007110; IG_1ike.                                      |              |              |
| DR       | InterPro; IPR003596; IG_v.   |              |              |
| DR       | Pfam; Pf00047; IG; 2.  |              |              |
| DR       | SMART; SM00409; IG; 2.   |              |              |
| DR       | SMART; SM00406; IGV; 2.  |              |              |
| DR       | PROSITE; PSS0835; IG LIKE; 2.                                      |              |              |
| FT       | NON TER  |              |              |
| FT       | 248  |              |              |
| FT       | 248  |              |              |
| SEQUENCE | 248 AA; 26634 MW; 7A3759BA3E570950 CRC64;                          |              |              |

```

Query Match      48.7%; Score 79.5; DB 2; Length 248;
Best Local Similarity 62.6%; Pred. No. 4,4e-50;
Matches 154; Conservative 29; Mismatches 52; Indels 11; Gaps 2;

QY      3 EVKLGSSGGLVDPGGRKLSCAASGTFSSFGCHWVRQAPKGLKLEWAVYSSSSSTIYY 62
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|:|

```



```

Db      2 DVKLVSGGGLVPGGSLKSLCATSGFTSDYYMMVVRQTEKRLWVAVISNDSSAAV 61
Qy      63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDVTVMYVYCARD-----YGA YMGQGTTVVSSG 117
Db      62 SDTVKGRFTISRDNPKNTLFLQMTSLRSEDVTVMYVYCARD-----YGA YMGQGTTVVSSG 121
Qy      118 GGGSGGGGGGGGGSDIELTQSPAIMSAPGERVMTCSASSV-----RYNMVFOQKSG 171
Db      122 GGGSGGGGGGGGGSDVLTQSPSLPLVSLGDQASISCRSQIIVHSNGNTYLEMYLQKPG 181
Qy      172 TSPKRWIYDTSKLSGVPAFAFSSGGSGTSLTITISMEADAAATYYCOQWSSNPLTFGAG 231
Db      182 QSPKLLIKVKNRFGVPRFSSGGSGTDTLKRISVEADLVGYCCFGQSHVPTFGSG 241
Qy      232 TKLELK 237
Db      242 TKLEIK 247

RESULT 5
065ZC8      PRELIMINARY;      PRT;      244 AA.
AC      065ZC8;
DT      25-OCT-2004 (TREMBLrel. 28, Created)
DT      25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE      25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE      Single-chain Fv (Fragment).
GN      Name=scfv;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97362799; PubMed=9219263;
RA      Kontermann R.E., Wing M.G., Winter G.;
RT      "Complement recruitment using bispecific antibodies.";
RL      Nat. Biotechnol. 15:629-631(1997).
DR      EMBL; Y13057; CAA73500.1; -
DR      InterPro; IPR003599; Ig.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003596; Ig_v.
DR      Pfam; PF00047; Ig_2.
DR      SMART; SM00409; IG_2.
DR      SMART; SM00406; IGV_2.
DR      PROSITE; PSS0835; IG_LIKE; 2.
FT      NON_TER 1
FT      NON_TER 244
SQ      SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;

Query Match      45.2%; Score 733; DB 2; Length 244;
Best Local Similarity 56.6%; Pred. No. 7.2e-46;
Matches 138; Conservative 45; Mismatches 53; Indels 8; Gaps 3;

```

```

RESULT 6
065ZL2      PRELIMINARY;      PRT;      487 AA.
AC      065ZL2;
DT      25-OCT-2004 (TREMBLrel. 28, Created)
DT      25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE      Fv/M4.
GN      Name=M4-IFN-<tau>;
OS      Mus sp.
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10095;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96272580; PubMed=8688499;
RA      Qi Y., Xiang J.;
RT      "A genetically engineered single-gene-encoded anti-TAC72 chimeric
RT      antibody secreted from myeloma cells.";
RL      Hum. Antibodies Hybridomas 6:161-166(1995).
DR      EMBL; S82493; AAB37424.2; -
DR      InterPro; IPR003599; Ig.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003597; Ig_c1.
DR      InterPro; IPR003598; Ig_c2.
DR      InterPro; IPR003006; IG_MHC.
DR      InterPro; IPR003596; IG_v.
DR      Pfam; PF07654; C1-set; 2.
DR      Pfam; PF00047; Ig_4.
DR      SMART; SM00409; IG_3.
DR      SMART; SM00407; IGC1; 2.
DR      SMART; SM00408; IGC2; 2.
DR      SMART; SM00406; IGV_2.
DR      PROSITE; PSS0835; IG_LIKE; 4.
DR      PROSITE; PSS0290; IG_MHC; UNKNOWN 1.
SQ      SEQUENCE 487 AA; 53578 MW; C7BAB69F30555504 CRC64;

Query Match      43.6%; Score 707; DB 2; Length 487;
Best Local Similarity 54.2%; Pred. No. 1.3e-43;
Matches 129; Conservative 46; Mismatches 59; Indels 4; Gaps 3;

```

RP SEQUENCE FROM N.A.  
RX MEDLINE=98170165; PubMed=9509426;  
RA Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,  
RA Yi K.S., Suh P.G., Ryu S.H., Chung H.K.,  
RT "Cloning and characterization of cDNAs encoding VH and VL of a  
RT monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and  
RT generation of a single-chain Fv molecule (scFv).";  
RL Mol. Cells 7:816-819 (1997).  
DR EMBL; U88067; AAB48044.1; -  
DR PIR; S19965; S19965.  
DR PIR; S19967; S19967.  
DR PIR; S19968; S19968.  
DR PIR; S26325; S26325.  
DR HSSP; P01607; 1BMW.  
DR SMART; SMO0406; IGV; 2.  
DR PROSITE; PSS0835; IG\_LIKE; 2.  
DR NON\_TER 1 1  
FT NON\_TER 241 241  
SQ SEQUENCE 241 AA; 26086 MW; 0276887248BE9C771 CRC64;

|                       |                 |               |           |            |
|-----------------------|-----------------|---------------|-----------|------------|
| Query Match           | 42.8%           | Score 695     | DB 2      | length 241 |
| Best Local Similarity | 55.7%           | Pred. No.     | 4,4e-43   |            |
| Matches 137           | Conservative 40 | Mismatches 55 | Indels 14 | Gaps 5     |

```

QY      3  EVKIQEBSGGGGLVQPGSGSRKLSCAASGFSSFGHMWVQAPEKGLIEWAVTISGGSSTIYY 62
Dd      1  QVTKIQSGSPELKRRGKGTWKISCKASGYFTFDYGMNMMWQAAGKGLKMMGMWINTYTGDEPTY 60

QY      63  ADIVYKGRFTISRDNPKNTLFLQMTSLNSEDTVMYICARD---YGAIVGGGTTVTVSSGG 118
Dd      61  ADPKGRFAPFLETSASFTAYLQINNLKNEDPATYFCARKDLLRFYDYGSGTIVTVSSGG 120

QY      119  GSGSGSGSGSGGSDIELTQSPAIMSAPGEHVMTTCSASSV-RYMNVPQKSGTSPRR- 176
Dd      121  GSGSGSGSGSGGSDIELTQSPSSLSASLGGKVTITCKASQDINKYIANYQHKPKGGRPSA 180

QY      177  ---WIYTSKLSGVPPARFSGSGSGTYSLTSSNEADAAITYYCOQSSNPLTFAGCT 233
Dd      181  HTLTHTY----IQGPISRFSGSGSGGRDYSFISINLEPEDIDATYYCLAH-DNLHTFGGCTK 235

QY      234  LELKRA 239
Dd      236  LELKRA 241

```

| QTTM2 | QTTM2   | PRELIMINARY                             | PRT | 243 AA. |
|-------|---|---|-----|---------|
| ID    | QTTM2   |   |     |         |
| AC    | QTTM2   |   |     |         |
| DT    | 01-OCT-2003   | (TREMBLrel. 25, Created)                |     |         |
| DT    | 01-OCT-2003   | (TREMBLrel. 25, Last sequence update)   |     |         |
| DT    | 01-MAR-2004   | (TREMBLrel. 26, Last annotation update) |     |         |
| DE    | ScFv 6H8 protein (Fragment).  |   |     |         |
| GN    | Name=scFv 6H8   |   |     |         |
| OS    | Mus musculus (Mouse)  |   |     |         |
| OC    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                                     |   |     |         |
| OC    | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.                                    |   |     |         |
| OX    | NCBI_TaxID=10090;   |   |     |         |
| RP    | [1]   |   |     |         |
| RP    | SEQUENCE FROM N.A.  |   |     |         |
| RC    | STRAIN=Balb/C;  |   |     |         |
| RX    | MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;  |   |     |         |
| RA    | Peter U.C., Ettekhatt P., Billiard P., Wallukat G., Hoebeke U.;                                       |   |     |         |
| RT    | "scFv single chain antibody variable fragment as inverse agonist for the beta-2 adrenergic receptor." |   |     |         |
| RL    | J. Biol. Chem. 278:36740-36747(2003).   |   |     |         |
| DR    | EMBL, AJ574851, CAFE00495.1; -  |   |     |         |
| DR    | HSSP, P01751, 1A6W.   |   |     |         |
| DR    | InterPro, IPR007110, IG_1like.  |   |     |         |
| DR    | InterPro, IPR003596, IG_v.  |   |     |         |
| DR    | SMART, SM00406, IGV; 2  |   |     |         |
| DR    | PROSITE, PSS0835, IG_LIKE; 2.   |   |     |         |
| NT    | NON_TER   |   |     |         |

|                       |                  |                    |           |             |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match           | 40.6%            | Score 658;         | DB 2;     | Length 243; |
| Best Local Similarity | 53.1%            | Pred. NO. 2.3e-40; |           |             |
| Matches 127;          | Conservative 43; | Mismatches 63;     | Indels 6; | Gaps 4;     |

Dy 1 EVYLQGGSGGLVLRPGASVKLSCKRASSGTTFTTYYMMHWKQRHHGGLEWICNITYPGSGITNY 60

QY 63 ADIVKGGFTTISNDNKNPTLFLOMTSLSESDITMYIYCARDYGA--YWGQGITIVIVSSGGGG 120

Db 61 DEKFKNGKILFVDTSSSTAYMHLSLASEDSAVYICARGGRGLDVGWCAGTTLLTIVSSGGGG 120

121 SGGGGSGGGSDIQMTOSSSSFVSICDVRVITTCASEDIYNNRLAWYOOKPGNAPRLIIS 180  
 121 SGGGGSGGGSDIETLSPALMSASPGKRVIMTCSSASSV-KRMNMFQDSSGISPKRMVY 179  
 121 SGGGGSGGGSDIETLSPALMSASPGKRVIMTCSSASSV-KRMNMFQDSSGISPKRMVY 179

DY  
180 TTSKLSGGVFAAFSGSSGISTSLTSLMEADAAITTCQ--MSSNPLTFAGAKLTBLA 237

Db  
181 GATSLFGLGVSPFSFGSSGCKDTLSTLSQTEDVAITYCQGYWSTR--TFGGGKTLEIK 237

Journal of  
Clinical  
Epidemiology

RESULT 9  
Q925S1  
PRELIMINARY:  
ID Q925S1  
PRT: 218 AA.

|    |             |    |   |
|----|-------------|----|---|
| AC | 09-03-95    | 1  | (Created)                               |
| DC | 01-DEC-2001 | 19 | (TREMBLrel. 19, Last sequence update)   |
| DT | 01-DEC-2001 | 19 | (TREMBLrel. 19, Last annotation update) |
| DT | 01-OCT-2003 | 25 | (TREMBLrel. 25, Last annotation update) |

OS Mus musculus (Mouse).  
 CC Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Metazoa;  
 Mammalia; Eutheria; Rodentia;  
 Sclerognathi; Muridae; Murinae; Mus.

|    |     |                    |
|----|-----|--------------------|
| RN | [1] | SEQUENCE FROM N.A. |
| RP |     |                    |
| RC |     | STRAIN=BALE/c;     |

RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,  
 RA Su C.: "Mechanism of exogenous nucleic acids and their precursors improving  
 RT the growth of *Chlamydomonas reinhardtii*," *Journal of Applied Microbiology*,  
 106, 102-109 (2008).

RL World J. Gastroenterol. 6:709-717 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC strain-Ba1B/C

RA Cui D., Zeng G., Yan X., Li X., Su C.;  
 RA "Cloning of mouse genes related to repairing  
 RT of the irradiated mice by treatment with the  
 RT the same strain".

DR SMART: SM004006: IGV: 1.  
DR HSSP: P01665: ION:  
DR EMBL: AF240168: AAK43733.1: -  
DR Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).  
RL

DR PROSITE; PS50835; 1G\_LIKE; 1.  
FT NON\_TER 218  
SQ SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

|                           |       |                |           |            |
|---------------------------|-------|----------------|-----------|------------|
| Query Match               | 37.5% | Score 608.5    | DB 2      | Length 218 |
| Best Local Similarity     | 55.5% | Pred. No. 8    | 7e-37     |            |
| Matches 121; Conservative | 33;   | Mismatches 55; | Indels 9; | Gaps 3;    |

Qy 1 MAEVLKDESGGGLVOPGSGRKLSCLMASGFTFSSFGHHVVRQAPBEKLEWVAIYSSGSSSTI 60  
Db 1 MAQVKLQDSGPELKKPGETVTRIS:::SCASGYTFTTAGMWMVQKPGFKGLMIWINTHSYVP 60

|    |    |   |        |               |     |
|----|----|---|--------|---------------|-----|
| Qy | 61 | YVADTVYGRFTISRDNKNTLEFLQMTSLRSEPTVMYCAR   | -YXG-- | -AYWGQGTIVTYS | 116 |
|    |    |   |        |               |     |
| Db | 61 | KAAEEFGREAFSLSTASTALVQLISDNKNEBTATYTCMRWD | DGGFA  | YWGQGTIVTYS   | 120 |
|    |    |   |        |               |     |

Oy 117 GGGSGGGGGSIDIELTSPAIMSASPGERVTMTCSASSV-----RYNWPFQOKSG 171  
|||||:::|||||:::|||

Df 121 GGGGSGGGGSSGGSDIVLTQSPASISLAVSLGQRATISGRAESVNIIGISFMMNQKRG 180

Oy 172 TSPKRWIYDTRSKTSGVPARFSGGSGCGSYSTLTSSME 209  
| | | | |  
Db 181 QPKLLIYAASKQGSVPAGLLASGDSFDLSINTYPME 218

RESULT 10

06PDB8 PRELIMINARY; PRT; 485 AA.

ID 06PDB8  
AC 06PDB8;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_Taxid=10090;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Colon;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
LA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein J.J., Ustin T.B., Toshitsuki S., Cantucci P., Prange C.,  
RA Roha S.S., Logguelano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Boeak S.A., McEwan P.J., McKernan K.T., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywicki M.I., Skalske U., Smallus D.E., Schnerch A., Schein J.E.,  
RT Jones S.J., Maiz M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Colon;  
RA Strausberg R.;  
RX Submitted (SEP-2003) to the EMBL/Genbank/DDbj databases.  
EMBL BC058814; AAH58814.1; --  
DR HSSP; P01861; IADQ.  
DR InterPro; IPRO03599; IG.  
DR InterPro; IPRO07110; IG\_1-like.  
DR InterPro; IPRO03597; IG\_CL.  
DR InterPro; IPRO03066; IG\_MHC.  
DR InterPro; IPRO03596; IG\_V.  
DR Pfam; PF07654; Cl-set; 2.  
DR SMART; SM00409; IG; 3.  
DR SMART; SM00407; IGc1; 3.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PSS0835; IG\_LIKE; 4.  
DR PROSITE; PSS0290; IG\_MHC; UNKNOWN\_2.  
KW Hypothetical protein.  
SQ SEQUENCE 485 AA; 52472 MW; 81236FF3AD821056 CRC64;

Query March 33.5%; Score 543; DB 2; Length 485;  
Beet Local Similarity 43.1%; Pred. No. 1.4e-31;  
Matches 140; Conservative 30; Mismatches 77; Indels 78; Gaps 13

Oy 3 EVKLQAESGGGLVOPGSRRLSCCAAGFTFFSFGGMVMVQAPEKGLEWVAIISSGSTITY 62  
:::|||||||:  
Db 20 DVQLVESGGGLVPPGSRRLSCCAAGFTFFSNNGMVMVQAPEKGLEWVAIISSSGITIFY 79  
:::|||||||:

Oy 63 ADIVAKGRFTTIRBNKNTLPLDMITSLRSEDTVMYICAR---DYGA--YMGQGTTVTYS 115

[illegible]

```

Query Match      33.4%; Score 542; DB 2; Length 473;
Best Local Similarity 51.7%; Pred. No. 1.6e-31;
Matches 120; Conservative 20; Mismatches 52; Indels 40; Gaps 5;

Oy 3 EYKLOESGGGLVOPGSGSKLSCAASGFTFSSFGMHVWQAPKGLWVAIYSSGSTIY 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 EYQLVESGGGLVQPGSGSKLSCAASGFTFSDYGMHVWQAPKGLWVAIYSSGSTIY 79
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Oy 63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDYVWYICARDYGA---YWGQGTIVTVSSG 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 80 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDYVWYICARDYGAELMRIDYWGQGTIVTVSSA- 138
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Oy 119 GSGSGGSGGSGGSDIELTQSPALMSASPERVYMTQSSSVRYVMWFOQKSTSTKRI 178
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 139 -----KTPPVSYPPLAPGCGDTGSSVTLCGLVKGYFPESVTVT----- 177
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Oy 179 YDTSKLSGV---PARFSGSGSGTSTSLTSSMEADATYVCCQWSSNPLT 227
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 178 WMSGSLSSSVHTFPALQSGLYTMSSSVTVPS-----STWPSQTVT 218
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
HVI6_MOUSE STANDARD; PRT; 136 AA.
AC P01783;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ig heavy chain V region MOPC 21 precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81234548; PubMed=6788376; DOI=10.1016/0092-8674(81)90089-1;
RA Botwell A.L.M., Paskind M., Reith M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
RN [2]
RP SEQUENCE OF 17-136.
RX MEDLINE=77100368; PubMed=401950;
RA Adecugbo K., Milstein C., Secher D.S.;
RT "Molecular analysis of spontaneous somatic mutants.";
RL Nature 265:299-304(1977).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL J00522; AADIS290.1; -
CC PIR E90809; GIM521.
CC PDB 1IGC; X-ray; H=-.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00835; IG_LIKE; 1.
KW 3D-structure; Direct protein sequencing; Immunoglobulin V region;
KW Signal.
FT SIGNAL 1 1
FT NON_TER <1 16
FT CHAIN 17 136 Ig heavy chain V region MOPC 21.
FT DOMAIN 115 119 D segment.
FT DOMAIN 120 136 JH4 segment.
FT DISULFID 38 112
FT CONFLICT 75 78 HYAD -> DYAH (in Ref. 2).

```

```

FT CONFLICT 89 90 DN -> ND (in Ref. 2).
FT CONFLICT 115 115 W -> H (in Ref. 2).
FT CONFLICT 120 120 Y -> W (in Ref. 2).
FT STRAND 19 23
FT STRAND 26 28
FT STRAND 30 31
FT STRAND 34 41
FT STRAND 45 47
FT STRAND 50 55
FT STRAND 61 67
FT TURN 69 70
FT STRAND 74 76
FT HELIX 78 80
FT STRAND 81 81
FT TURN 82 83
FT STRAND 84 89
FT TURN 90 93
FT STRAND 94 99
FT HELIX 104 106
FT STRAND 108 114
FT TURN 118 119
FT STRAND 125 126
FT STRAND 130 134
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15071 MW; 2276398DBDF7016 CRC64;

Query Match      33.0%; Score 535; DB 1; Length 136;
Best Local Similarity 87.5%; Pred. No. 1.2e-31;
Matches 105; Conservative 5; Mismatches 4; Indels 6; Gaps 2;

Oy 3 EYKLOESGGGLVOPGSGSKLSCAASGFTFSSFGMHVWQAPKGLWVAIYSSGSTIY 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 17 DVALVESGGGLVQPGSGSKLSCAASGFTFSSFGMHVWQAPKGLWVAIYSSGSTIHY 76
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Oy 63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDYVWYICAR---DYGA--YWGQGTIVTVSS 116
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 77 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDYVWYICARMGNYPYAMDYWGQGTIVTVSS 136
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
KV6F_MOUSE STANDARD; PRT; 107 AA.
AC P04940;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-VI region NQ2-17.4.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83721467; PubMed=6877353;
RA Kaartinen M., Griffiths G.M., Matham A.F., Milstein C.;
RT "mRNA sequences define an unusually restricted IgG response to 2-
RT phenylloxazalone and its early diversification.";
RL Nature 304:320-324(1983).
CC -----
CC -I- FUNCTION: Anti-2-phenyl oxazalone (PROX) Antibody.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL K00735; AAA38680.1; -
CC HSSP: P01679; 2FBJ.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.

```

DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS0835; IG LIKE; 1.  
 KW Hybridoma; Immunoglobulin V region.  
 FT DOMAIN 1 23 Framework-1.  
 FT DOMAIN 24 33 Complementarity-determining-1.  
 FT DOMAIN 34 48 Framework-2.  
 FT DOMAIN 49 55 Complementarity-determining-2.  
 FT DOMAIN 56 87 Framework-3.  
 FT DOMAIN 88 96 Complementarity-determining-3.  
 FT DOMAIN 97 106 Framework-4.  
 FT DISULFID 23 87 By similarity.  
 FT NON\_TER 107 107  
 SQ SEQUENCE 107 AA; 11561 MW; 6F694284ECFA68B6 CRC64;

Query Match 31.3%; Score 507; DB 1; Length 107;  
 Best Local Similarity 90.6%; Pred. No. 1,1e-29;  
 Matches 96; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 133 IELTOSPAIMASPGERVMTCSASSSVRYMFMFOOKSGTSPKRWIYDTSKLSGVDPARF 192  
 DB 2 IVLTOSPAIMASPGQKVMTCSASSSVSYMHVYQOKSGTSPKRWIYDTSKLSGVDPARF 61  
 QY 193 SGGSGSTSYSLTTSMEADPAATYYCOQSSNPLTFGAGTKLEIKR 238  
 DB 62 SGGSGSTSYSLTTSMQAEDPAATYYCOQSSNPLTFGAGTKLEIKR 107

RESULT 14  
 KV6I\_MOUSE STANDARD; PRT; 107 AA.  
 AC P04942;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG kappa chain V-VI region HQ5-61.1.2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83271467; PubMed=6877353;  
 RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
 RT "mRNA sequences define an unusually restricted IGH response to 2-phenylloxazalone and its early diversification."  
 RL Nature 304:320-324(1983).  
 CC -1- FUNCTION: Anti-2-phenyl oxazalone (PHOX) Antibody.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; K00739; AAA38684.1; -  
 DR HSSP; P01679; 2FBJ  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003596; IG\_v.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS0835; IG LIKE; 1.  
 KW Hybridoma; Immunoglobulin V region.  
 FT DOMAIN 1 23 Framework-1.  
 FT DOMAIN 24 33 Complementarity-determining-1.  
 FT DOMAIN 34 48 Framework-2.  
 FT DOMAIN 49 55 Complementarity-determining-2.  
 FT DOMAIN 56 87 Framework-3.  
 FT DOMAIN 88 96 Complementarity-determining-3.  
 FT DOMAIN 97 106 Framework-4.  
 FT DISULFID 23 87 By similarity.  
 FT NON\_TER 107 107

SQ SEQUENCE 107 AA; 11605 MW; CA6C4284EFCB550 CRC64;

Query Match 31.2%; Score 506; DB 1; Length 107;  
 Best Local Similarity 90.6%; Pred. No. 1.3e-29;  
 Matches 96; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 133 IELTOSPAIMASPGERVMTCSASSSVRYMFMFOOKSGTSPKRWIYDTSKLSGVDPARF 192  
 DB 2 IVLTOSPAIMASPGQKVMTCSASSSVSYMHVYQOKSGTSPKRWIYDTSKLSGVDPARF 61  
 QY 193 SGGSGSTSYSLTTSMEADPAATYYCOQSSNPLTFGAGTKLEIKR 238  
 DB 62 SGGSGSTSYSLTTSMQAEDPAATYYCOQSSNPLTFGAGTKLEIKR 107

RESULT 15  
 KV6I\_MOUSE STANDARD; PRT; 107 AA.  
 AC P04943;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG kappa chain V-VI region HQ6-8.3.1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83271467; PubMed=6877353;  
 RA Kaartinen M., Griffiths G.M., Markham A.F., Milstein C.;  
 RT "mRNA sequences define an unusually restricted IGH response to 2-phenylloxazalone and its early diversification."  
 RL Nature 304:320-324(1983).  
 CC -1- FUNCTION: Anti-2-phenyl oxazalone (PHOX) Antibody.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; K00740; AAA38685.1; -  
 DR HSSP; P01679; 2FBJ  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003596; IG\_v.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS0835; IG LIKE; 1.  
 KW Hybridoma; Immunoglobulin V region.  
 FT DOMAIN 1 23 Framework-1.  
 FT DOMAIN 24 33 Complementarity-determining-1.  
 FT DOMAIN 34 48 Framework-2.  
 FT DOMAIN 49 55 Complementarity-determining-2.  
 FT DOMAIN 56 87 Framework-3.  
 FT DOMAIN 88 96 Framework-4.  
 FT DISULFID 23 87 By similarity.  
 FT NON\_TER 107 107  
 SQ SEQUENCE 107 AA; 11572 MW; 6F694284EFC0CB6 CRC64;

Query Match 30.9%; Score 502; DB 1; Length 107;  
 Best Local Similarity 89.6%; Pred. No. 2.5e-29;  
 Matches 95; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 133 IELTOSPAIMASPGERVMTCSASSSVRYMFMFOOKSGTSPKRWIYDTSKLSGVDPARF 192  
 DB 2 IVLTOSPAIMASPGQKVMTCSASSSVSYMHVYQOKSGTSPKRWIYDTSKLSGVDPARF 61  
 QY 193 SGGSGSTSYSLTTSMEADPAATYYCOQSSNPLTFGAGTKLEIKR 238

Db 62 SSGSATSYSLITSMQAEADATYCCQMSNPITFGAGTKLELR 107

Search completed: June 23, 2005, 05:52:29  
Job time : 86 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 05:50:26 ; Search time 77 Seconds  
(without alignments)  
1587.224 Million cell updates/sec

Title: US-09-403-882a-2

Perfect score: 1622  
Sequence: 1 MAEVLKQSGGLVQPGGSR.....VLTITLITLIMWKKER 316

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Geneseq\_16Dec04:\*

1: geneseqp1980a:\*\n2: geneseqp1980a:\*\n3: geneseqp2000a:\*\n4: geneseqp2001a:\*\n5: geneseqp2002a:\*\n6: geneseqp2003a:\*\n7: geneseqp2003ba:\*\n8: geneseqp2004a:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID        | Description         |
|------------|--------|-------------|--------|-----------|---------------------|
| 1          | 1622   | 100.0       | 316    | AAV32086  | Aay32086 Single ch  |
| 2          | 1319   | 81.3        | 255    | AAAB1398  | Aab1398 E. coli e   |
| 3          | 1319   | 81.3        | 255    | AAAB7419  | Aab74199 PeLB-ecFv  |
| 4          | 1319   | 81.3        | 255    | AAAB7076  | Aab707620 Expressio |
| 5          | 1319   | 81.3        | 255    | AAAB7202  | Aay72020 E. caroto  |
| 6          | 1237   | 76.3        | 241    | ABR62010  | Abc62010 Single-ch  |
| 7          | 1227   | 75.6        | 241    | AAAB48925 | Aam48925 scFv anti  |
| 8          | 1194   | 73.6        | 236    | AAAR32842 | Aar32842 VH NQ10/1  |
| 9          | 1185   | 73.1        | 240    | AAAR68613 | Aar68613 Single ch  |
| 10         | 1097.5 | 67.7        | 223    | AAAR32841 | Aar32841 VH NQ10/1  |
| 11         | 1008   | 62.1        | 580    | AAAB90217 | Aab90217 Bispecific |
| 12         | 1001   | 61.7        | 556    | AAAB90218 | Aab90218 Bispecific |
| 13         | 976.5  | 60.2        | 237    | AAAB38657 | Aab38657 Mouse GI   |
| 14         | 970.5  | 59.8        | 268    | AAAY44973 | Aay44973 Recombina  |
| 15         | 967.5  | 59.6        | 268    | AAAY44972 | Aay44972 Recombina  |
| 16         | 964.5  | 59.5        | 284    | AAAB95569 | Aab95569 Intracell  |
| 17         | 959    | 59.1        | 281    | AAAB95569 | Aab95569 Intracell  |
| 18         | 954    | 58.8        | 280    | AAAB95569 | Aab95569 Intracell  |
| 19         | 939    | 57.9        | 234    | AAAB42055 | Ab42055 Newcaetle   |
| 20         | 938.5  | 57.9        | 235    | AAAB42055 | Ab42055 Newcaetle   |
| 21         | 937.5  | 57.8        | 272    | AAAB42055 | Ab42055 Newcaetle   |
| 22         | 937    | 57.8        | 272    | AAAB42055 | Ab42055 Newcaetle   |
| 23         | 935.5  | 57.7        | 258    | AAAB90221 | Aab90221 Anti-B7.1  |
| 24         | 934.5  | 57.6        | 234    | AAAB20442 | Aab20442 Anti-FIX   |
| 25         | 928.5  | 57.2        | 282    | AAAB5715  | Aab5715 Anti-huma   |

|    |       |      |     |   |           |                    |
|----|-------|------|-----|---|-----------|--------------------|
| 26 | 927   | 57.2 | 291 | 8 | ADN06992  | Adn06992 Human EFG |
| 27 | 926   | 57.1 | 287 | 8 | ADN06988  | Adn06988 Human EFG |
| 28 | 921.5 | 56.8 | 244 | 8 | ADQ75290  | Adq75290 Immunoglo |
| 29 | 919   | 56.7 | 553 | 2 | AAW11508  | Aaw11508 Single ch |
| 30 | 919   | 56.7 | 553 | 2 | AAW73223  | Aaw73223 H22-anti- |
| 31 | 919   | 56.7 | 553 | 4 | AAAB1960  | Aab1960 Bispecific |
| 32 | 919   | 56.7 | 553 | 4 | AAAB8545  | Aab8545 Bispecific |
| 33 | 917.5 | 56.6 | 269 | 8 | ADSI17525 | Adsi17525 Anti-act |
| 34 | 913   | 56.3 | 242 | 6 | ABR62132  | Abc62132 Single ch |
| 35 | 913   | 56.3 | 242 | 6 | ADT91209  | Adt91209 Single ch |
| 36 | 912   | 56.2 | 242 | 8 | ADQ75289  | Adq75289 Immunoglo |
| 37 | 907   | 55.9 | 353 | 2 | AAAY06273 | Aay06273 Anti-Fc a |
| 38 | 906.5 | 55.9 | 237 | 5 | ABP46093  | Abp46093 Human Bly |
| 39 | 906.5 | 55.9 | 237 | 7 | ADG96920  | Adg96920 Single ch |
| 40 | 906.5 | 55.9 | 253 | 2 | AAAB90225 | Aab90225 Anti-B7.1 |
| 41 | 903.5 | 55.7 | 237 | 5 | ABP46009  | Abp46009 Human Bly |
| 42 | 903.5 | 55.7 | 237 | 7 | ADG96836  | Adg96836 Single ch |
| 43 | 900.5 | 55.5 | 237 | 5 | ABP46103  | Abp46103 Human Bly |
| 44 | 900.5 | 55.5 | 237 | 5 | ABP45994  | Abp45994 Human Bly |
| 45 | 900.5 | 55.5 | 237 | 7 | ADG96930  | Adg96930 Single ch |

## ALIGNMENTS

|          |   |                            |
|----------|---|----------------------------|
| RESULT 1 | AAV32086  | standard; protein; 316 AA. |
| ID       | AAV32086  |                            |
| XX       | AAV32086;   |                            |
| AC       |   |                            |
| DT       | 17-JAN-2000 (first entry)   |                            |
| XX       |   |                            |
| DE       | Single chain antibody used in probe detection.                              |                            |
| XX       |   |                            |
| KM       | Single chain antibody; scab; sfv; spectroscopic probe.                      |                            |
| XX       |   |                            |
| OS       | Unidentified  |                            |
| XX       |   |                            |
| PD       | MO9951986-A1.   |                            |
| XX       | 14-OCT-1998   |                            |
| XX       |   |                            |
| PF       | 08-APR-1999; 99WC-US007847.   |                            |
| XX       |   |                            |
| PR       | 08-APR-1998; 98US-0081118P.   |                            |
| PR       | 09-APR-1998; 98US-0081340P.   |                            |
| XX       |   |                            |
| PA       | (REGC) UNIV CALIFORNIA.   |                            |
| XX       |   |                            |
| PI       | Farinas J;  |                            |
| XX       |   |                            |
| DR       | WFI; 1999-611066/52.  |                            |
| XX       | N-PDB; AAZ20266.  |                            |
| PT       | Localizing probes to specific sites in cells that express single-chain      |                            |
| PT       | antibody reactive with probe-11ligand conjugate, particularly for detecting |                            |
| PT       | post-translational modification and its modulators.                         |                            |
| XX       |   |                            |
| PS       | Claim 5; Page 47-48; 69pp; English.   |                            |
| XX       |   |                            |
| CC       | The present sequence represents a single chain antibody (scab) that has 2   |                            |
| CC       | c-myc epitopes. The scab is used in claimed methods for localizing a        |                            |
| CC       | probe and for detecting a post-translational activity in a cell, and is     |                            |
| CC       | expressed by a claimed transgenic non-human animal. The invention           |                            |
| CC       | provides methods and reagents for targeting probes to selected cellular     |                            |
| CC       | locations through the expression of specific binding partners (such as      |                            |
| CC       | scabs) within a cell, and for creating assays for post-translational        |                            |
| CC       | activities. The invention allows the monitoring of the location of such     |                            |
| CC       | intracellular, specific binding partners over time and in response to       |                            |
| CC       | stimuli, such as test chemicals. Spectroscopic probes can be used to        |                            |
| CC       | screen a test chemical for activity. The invention also includes cells      |                            |
| CC       | and transgenic organisms comprising the intracellular specific binding      |                            |



CC partner, where the specific binding partner can bind with the  
CC spectroscopic probe/ligand conjugate. scabs can be expressed within the  
CC cell and can be designed to bind a wide variety of spectroscopic probes,  
CC including small molecules that have better (and more diverse)  
CC spectroscopic properties than green fluorescent protein

XX Sequence 316 AA;

Query Match 100.0%; Score 1622; DB 2; Length 316;  
Best Local Similarity 100.0%; Pred. No. 2.8e-108;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAEVKLOESGGGLVOPGSRKLSCAASGFTFSSFGMHVWROAPEKLEWAVAYISSGSSSTI 60
DB 1 MAEVKLOESGGGLVOPGSRKLSCAASGFTFSSFGMHVWROAPEKLEWAVAYISSGSSSTI 60
QY 61 YYADTVKGRFTISRDNPKNTLFIQMTSLRSEDTVMYTCARDYGAWGGGTTVSSGGGG 120
DB 61 YYADTVKGRFTISRDNPKNTLFIQMTSLRSEDTVMYTCARDYGAWGGGTTVSSGGGG 120
QY 121 SGGGGSGGGSDIELTQSPALMSASPGERVMTCSASSSVRYMNFQOKSGTSPKRWIYD 180
DB 121 SGGGGSGGGSDIELTQSPALMSASPGERVMTCSASSSVRYMNFQOKSGTSPKRWIYD 180
QY 181 TSKLSSGVPARFSGSGSTSYSLTSSMEABDAATYYCOQWSSNPLTFGAGTKLELKRAA 240
DB 181 TSKLSSGVPARFSGSGSTSYSLTSSMEABDAATYYCOQWSSNPLTFGAGTKLELKRAA 240
QY 241 AEOKLISEEDLNGAVDEOKLISEEDLNAGODTQEVIVPHSLPFVVVYSAIILVLT 300
DB 241 AEOKLISEEDLNGAVDEOKLISEEDLNAGODTQEVIVPHSLPFVVVYSAIILVLT 300
QY 301 IISLIILIMLWQKKR 316
DB 301 IISLIILIMLWQKKR 316
```

RESULT 2  
AAB11398  
ID AAB11398 standard; protein; 255 AA.

XX AAB11398;

DT 22-FEB-2001 (first entry)

DE E. coli expression plasmid pUBS520-ScFvOX encoded protein.

KW Eukaryotic protein; protease; interferon; antibody; hormone;  
disulfide bridge.

OS Escherichia coli.

OS Synthetic.

PN EPI048732-A1.

PD 02-NOV-2000.

PF 26-APR-1999; 99EP-00107412.

PR 26-APR-1999; 99EP-00107412.

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

DR N-PSDB; AAC66074.

XX MPI; 2000-674185/66.

PT Preparation of water-soluble eukaryotic polypeptides with disulfide  
bridges e.g. rPA, comprises cultivation of prokaryotic cells in the  
presence of arginine or amide compound.

PS Example 6; Page 22-23; 40pp; German.

XX This invention describes a novel preparation of a water-soluble,

CC naturally occurring eukaryotic polypeptide containing two or more  
CC cysteine units bound via a disulfide bridge which comprises cultivation  
CC of prokaryotic cells in the presence of arginine or an amide compound.  
CC The method is useful for the preparation of eukaryotic proteins e.g.  
CC proteases, interferons, protein hormones, antibodies or antibody  
CC fragments (e.g. a single chain Fv fragment that binds to thyroid  
CC stimulating hormone). It is especially useful for preparing proteins with  
CC more than five disulfide bridges, e.g. recombinant plasmidogen activator  
CC (rPA). The technique is simple and does not require in vitro after-  
CC treatment, such as the removal of inclusion bodies, reduction or  
CC naturalization

XX Sequence 255 AA;

Query Match 81.3%; Score 1319; DB 3; Length 255;  
Best Local Similarity 99.2%; Pred. No. 1.2e-86;  
Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MAEVKLOESGGGLVOPGSRKLSCAASGFTFSSFGMHVWROAPEKLEWAVAYISSGSSSTI 60
DB 1 MAEVKLOESGGGLVOPGSRKLSCAASGFTFSSFGMHVWROAPEKLEWAVAYISSGSSSTI 60
QY 61 YYADTVKGRFTISRDNPKNTLFIQMTSLRSEDTVMYTCARDYGAWGGGTTVSSGGGG 120
DB 61 YYADTVKGRFTISRDNPKNTLFIQMTSLRSEDTVMYTCARDYGAWGGGTTVSSGGGG 120
QY 121 SGGGGSGGGSDIELTQSPALMSASPGERVMTCSASSSVRYMNFQOKSGTSPKRWIYD 180
DB 121 SGGGGSGGGSDIELTQSPALMSASPGERVMTCSASSSVRYMNFQOKSGTSPKRWIYD 180
QY 181 TSKLSSGVPARFSGSGSTSYSLTSSMEABDAATYYCOQWSSNPLTFGAGTKLELKRAA 240
DB 181 TSKLSSGVPARFSGSGSTSYSLTSSMEABDAATYYCOQWSSNPLTFGAGTKLELKRAA 240
QY 241 AEOKLISEEDLNGA 254
DB 241 AEOKLISEEDLNGA 254
```

RESULT 3  
AAB74199  
ID AAB74199 standard; protein; 255 AA.

XX AAB74199;

DT 29-MAY-2001 (first entry)

DE PelB-scFvOxazol fusion protein.

KW Molecular chaperone; PelB signal sequence; scFvOxazol.

OS Unidentified.

PN EPI077262-A1.

PD 21-FEB-2001.

PF 24-JUL-2000; 2000EP-00115839.

PR 29-JUL-1999; 99EP-00114811.

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

DR MPI; 2001-246712/26.

DR N-PSDB; AAF77806.

PT Producing naturally folded eukaryotic proteins e.g. antibodies,  
interferon, hormones or proteases that contain two or several cysteines  
linked by disulfide bridges comprises co-expression of a molecular  
chaperone.

XX

PS Disclosure; Page 19; 35pp; English.

XX The present invention relates to a method for production of a naturally  
 CC folded eukaryotic protein containing two or more cysteines linked by  
 CC disulfide bridges. The method comprises co-expression and secretion into  
 CC the periplasm of a molecular chaperone via an expression and secretion into  
 CC the periplasm of a molecular chaperone also encodes a signal sequence.  
 CC The method is useful for producing a naturally folded eukaryotic protein  
 CC such as an antibody, antibody fragment, interferon, protein hormone or a  
 CC process containing two or several cysteines linked by disulfide bridges.  
 CC The present sequence is a fusion protein composed of the pelb signal  
 CC sequence and ScFvOxazoln. This sequence was used in the method of the  
 CC present invention

XX Sequence 255 AA;

Query Match 81.3%; Score 1319; DB 4; Length 255;  
 Best Local Similarity 99.2%; Pred. No. 1.2e-86;  
 Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MAEVKIQESGGGLVOPGSRKLSCAASGFTFSFGHMWRQAPKGLKEMVAVYISSGSSSTI 60  
 1 MAEVKIQESGGGLVOPGSRKLSCAASGFTFSFGHMWRQAPKGLKEMVAVYISSGSSSTI 60  
 61 YVADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAVYGQGTIVTVSSGGGG 120  
 61 YVADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAVYGQGTIVTVSSGGGG 120  
 121 SGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSSVRYNMWFOQKSGTSPKMIYD 180  
 121 SGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSSVRYNMWFOQKSGTSPKMIYD 180  
 181 TSKLSSGVPAFPSSGGSGTSTLITSSMEADPAATYYCOQMSNPITFGAGTLELKRRA 240  
 181 TSKLSSGVPAFPSSGGSGTSTLITSSMEADPAATYYCOQMSNPITFGAGTLELKRRA 240  
 241 AEOKLISEEDLNGA 254  
 241 AEOKLISEEDLNGA 254

RESULT 4  
 AAB70769 standard; protein; 255 AA.

XX AAB70769;

DT 18-MAY-2001 (first entry)

DE Expression plasmid pUBS520-ScFvOx protein.

KW Chaperone protein; periplasm; antibody production; protein production;  
 KW Interferon production; protease production.

OS Escherichia coli.  
 OS Synthetic.

PN EP1077263-A1.

PD 21-FEB-2001.

PF 29-JUL-1999; 99EP-00114811.

PR 29-JUL-1999; 99EP-00114811.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX WPI; 2001-204356/21.

DR N-PSDB; AAF61193.

PT Preparation of naturally folded eukaryotic proteins, e.g. antibodies, by  
 PT simultaneous expression of a chaperone protein, allows simple recovery  
 PT from periplasm or medium.

XX Disclosure; Page 20-21; 36pp; German.

XX This invention describes a novel method for preparing a naturally folded  
 CC eukaryotic polypeptide (I) that contains two or more disulfide-bridged  
 CC Cys residues by culturing prokaryotic cells that contain an expression  
 CC vector for (I) including a prokaryotic signal sequence at its N-terminus  
 CC and a nucleic acid (II) that secretes a chaperone protein (III) into the  
 CC periplasm. (I) is secreted into the periplasm or medium; the signal  
 CC peptide is then cleaved and (I) isolated from the periplasm or medium.  
 CC The method is used for production of antibody, interferon, protein  
 CC hormone or protease. Expression of (III) increases the yield of (I). The  
 CC method is simple and eliminates time-consuming in vitro processing  
 CC operations such as dissolution of inclusion bodies, reduction and  
 CC refolding. (III) protects (I) against agglomeration and promotes their  
 CC natural conformation

XX Sequence 255 AA;

Query Match 81.3%; Score 1319; DB 4; Length 255;  
 Best Local Similarity 99.2%; Pred. No. 1.2e-86;  
 Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MAEVKIQESGGGLVOPGSRKLSCAASGFTFSFGHMWRQAPKGLKEMVAVYISSGSSSTI 60  
 1 MAEVKIQESGGGLVOPGSRKLSCAASGFTFSFGHMWRQAPKGLKEMVAVYISSGSSSTI 60  
 61 YVADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAVYGQGTIVTVSSGGGG 120  
 61 YVADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAVYGQGTIVTVSSGGGG 120  
 121 SGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSSVRYNMWFOQKSGTSPKMIYD 180  
 121 SGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSSVRYNMWFOQKSGTSPKMIYD 180  
 181 TSKLSSGVPAFPSSGGSGTSTLITSSMEADPAATYYCOQMSNPITFGAGTLELKRRA 240  
 181 TSKLSSGVPAFPSSGGSGTSTLITSSMEADPAATYYCOQMSNPITFGAGTLELKRRA 240  
 241 AEOKLISEEDLNGA 254  
 241 AEOKLISEEDLNGA 254

RESULT 5  
 AAY72020 standard; protein; 255 AA.

XX AAY72020;

DT 28-MAR-2001 (first entry)

DE E. carotovora pelb-scFvOx fusion protein encoded by pUBS520-pITN-scFvOx.

KW Secreted protein; chaperone; interferon; protease; hormone;  
 KW naturally folded protein; lac promoter; DnaJ; heat shock protein; HSP;  
 KW pectate lyase B; pelB; hapten; single-chain Fv-fragment Oxazoln;  
 KW ScFvOxazoln; fusion protein; thyroid stimulating hormone; TSH.

XX Pectobacterium carotovorum.

OS Undifferentiated.

OS Chimeric.

PN EP1054063-A2.

PD 22-NOV-2000.

PF 19-APR-2000; 2000EP-00108505.

PR 26-APR-1999; 99EP-00107412.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Ambrosius D, Rudolph R, Schaeffner J, Schwarz E;  
XX  
XX WPI; 2001-033777/05.  
DR N-PSDB; AAD02212.  
XX  
XX  
XX Producing water-soluble, naturally folded, and secreted eukaryotic  
PT polypeptide, involves culturing prokaryotic cells containing an  
PT expression vector encoding the polypeptide in the presence of arginine or  
PT a specific compound.

Example 6; Page 22-23; 35pp; English.

XX The patent discloses a method for the production of a water-soluble,  
CC naturally folded and secreted eukaryotic proteins in prokaryotic cells.  
CC The method involves culturing the prokaryotic cells, containing an  
CC expression vector encoding the desired protein and the prokaryotic signal  
CC sequence, in the presence of an additive, e.g., L-arginine. The signal  
CC sequence mediates the secretion of the desired protein into the  
CC periplasm, where folding of the protein takes place. The prokaryotic cell  
CC also contains an expression vector encoding a molecular chaperone, e.g.,  
CC DnaK and heat shock protein 25 (HSP25). The simultaneous co-  
CC overexpression and co-secretion of molecular chaperones in the periplasm  
CC improves the yield of functionally folded protein. The above method  
CC recombinantly produces a high yield of eukaryotic secreted proteins in  
CC prokaryotes. The method is useful for producing eukaryotic proteins such  
CC as an antibody, antibody fragment, interferon, protein hormone or a  
CC protease. The present sequence is an Erwinia carotovora pectate lyase B  
CC (PelB) signal sequence-scFvOX fusion protein encoded by PUBS520-PIN-  
CC scFvOX. The plasmid, pUBS520-PIN-scFvOX, also comprises the lac promoter.  
CC The single-chain Fv-fragment, which is directed against the hapten  
CC oxazolone (SCFOxazolone), is an antibody fragment against thyroid  
CC stimulating hormone (TSH). The co-expression of scFvOX which has no  
CC chaperone properties is used as a negative control

Sequence 255 AA;

Query Match 81.3%; Score 1319; DB 4; Length 255;

Best Local Similarity 99.2%; Pred. No. 1.2e-86;

Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAEVLKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHWROAPEKGLWVAIYSSGSTI 60  
DB 1 MAEVLKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHWROAPEKGLWVAIYSSGSTI 60  
QY 61 YYADTVKGRFTTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAWGQGTTVVSSGGGG 120  
DB 61 YYADTVKGRFTTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAWGQGTTVVSSGGGG 120  
QY 121 SGGGSGGGGSDIELTQSPAIMSASPERVMTCSASSSVRYMNMFOQSGTSPKRWIYD 180  
DB 121 SGGGSGGGGSDIELTQSPAIMSASPERVMTCSASSSVRYMNMFOQSGTSPKRWIYD 180  
QY 181 TSKLSSGVPARFSGSGSGTYSGLTSSMEADPAATYYCQWSSNPLTFGAGTKLEKRAA 240  
DB 181 TSKLSSGVPARFSGSGSGTYSGLTSSMEADPAATYYCQWSSNPLTFGAGTKLEKRAA 240  
QY 241 AEQKLISEEDLNGA 254  
DB 241 AEQKLISEEDLNGA 254

RESULT 6

ABR62010 standard; protein; 241 AA.

ABR62010;

03-OCT-2003 (first entry)

Single-chain (scFv) antibody.

Transgenic; biomolecule; ATP; ADP; cyostatic; virucide; gene therapy;  
antibiose therapy; scFv; antibody.

XX Synthetic.  
OS  
XX  
XX EPI321524-A1.  
PN  
XX  
XX 25-JUN-2003.  
PD  
XX  
XX 19-DEC-2001; 2001EP-00130319.  
PF  
XX  
XX 19-DEC-2001; 2001EP-00130319.  
PR  
XX  
XX (DUEP/) DUEP K.  
PA  
XX  
XX Mahn A, Hanke S, Peters D;  
PI  
XX  
XX WPI; 2003-543829/52.  
DR  
XX  
XX N-PSDB; ACC84876.

XX Increasing the content of transgene-coded biomolecules in a plant or  
PT animal, useful for producing proteins for diagnosing, preventing and/or  
PT treating viral diseases and cancer, comprises changing the distribution  
PT of ATP and/or ADP.

Example 2; Fig 2; 18pp; English.

XX The invention relates to increasing the content of one or more transgene-  
CC coded biomolecules in an organism and involves changing the distribution  
CC of ATP and/or ADP in cells of the organism. The yield of transgenic  
CC molecules in host cells is often insufficient for industrial production.  
CC The method increases the yield of transgenic molecules in animal and  
CC plant host cells, therefore facilitating their production on an  
CC industrial scale. The proteins produced by the method are useful for  
CC diagnosing, preventing and/or treating viral diseases and cancer. The  
CC present sequence represents a single-chain (scFv) antibody, used to  
CC exemplify the increase in the expression of scFv antibodies in transgenic  
CC potato tubers

Sequence 241 AA;

Query Match 76.3%; Score 1237; DB 6; Length 241;

Best Local Similarity 97.9%; Pred. No. 8.5e-81;

Matches 235; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 AEVLKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHWROAPEKGLWVAIYSSGSTI 61  
DB 2 AEVLKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHWROAPEKGLWVAIYSSGSTI 61  
QY 62 YADTVKGRFTTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAWGQGTTVVSSGGGG 121  
DB 62 YADTVKGRFTTISRDNPKNTLFLQMTSLRSEDTVMYYCARDYGAWGQGTTVVSSGGGG 121  
QY 122 GGGGSGGGGSDIELTQSPAIMSASPERVMTCSASSSVRYMNMFOQSGTSPKRWIYD 181  
DB 122 GGGGSGGGGSDIELTQSPAIMSASPERVMTCSASSSVRYMNMFOQSGTSPKRWIYD 181  
QY 182 SKLSSGVPARFSGSGSGTYSGLTSSMEADPAATYYCQWSSNPLTFGAGTKLEKRAA 241  
DB 182 SKLSSGVPARFSGSGSGTYSGLTSSMEADPAATYYCQWSSNPLTFGAGTKLEKRAA 241

RESULT 7

AAM48925 standard; protein; 241 AA.

AAM48925;

03-MAY-2002 (first entry)

scFv antibody.

Ketone binding protein; oxazole; pathogen resistance; virucide;  
fungicide; antibacterial; scFv antibody.

OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 137 /note= "encoded by CAG"  
FT Misc-difference 201 /note= "encoded by TAC"  
FT  
XX WO200204020-A2.  
XX  
XX  
XX PD 17-JAN-2002.  
XX  
XX PF 17-MAY-2001; 2001WO-DE001916.  
XX  
XX PR 12-JUL-2000; 2000DE-01033750.  
XX  
XX PA (MPBC-) MPB COLOGNE GMBH.  
XX  
XX PI Duering K, Brinkmann O;  
XX  
XX DR MPI; 2002-154868/20.  
XX DR N-PSDB; AAK98639.  
XX  
XX PT Imparting pathogen resistance to plants and animals, comprises using a  
XX PT polypeptide that binds to an oxazole-derived ketone, optionally expressed  
XX PT from nucleic acid.  
XX  
XX PS Example 3; Fig 1; 20pp; German.  
XX  
XX CC The present invention relates to the use of a polypeptide that binds a  
XX CC ketone derived from oxazole for generating pathogen resistance in an  
XX CC organism. This can be used to impart resistance to pathogens (bacteria,  
XX CC fungi or viruses) to a wide range of plants (e.g. cereals, sugar beet,  
XX CC tobacco etc.), humans, farm animals and pets. Exemplified are activities  
XX CC against Phytophthora infestans and Erwinia carotovora in potatoes and  
XX CC Stephylococcus aureus in mice. The present sequence is a 8c-Fv antibody,  
XX CC which is capable of binding 2-phenyloxazol-5-one  
XX  
SQ Sequence 241 AA;  
Query Match 75.6%; Score 1227; DB 5; Length 241;  
Best Local Similarity 97.1%; Pred. No. 4.5e-80;  
Matches 234; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MAEVLQESGGGLVOPGSGSRKLSCAASGFTFSSFGMHWRQAPKGLQEMVAIYSSGSSSTI 60  
DB 1 MADVLVBSGGGLVOPGSGSRKLSCAASGFTFSSFGMHWRQAPKGLQEMVAIYSSGSSSTI 60  
QY 61 YYADTVKGRFTISRDNPKNTLFLQMTSLRSEDVTMYTCARDYGAWGCGTTVYSSGGG 120  
DB 61 YYADTVKGRFTISRDNPKNTLFLQMTSLRSEDVTMYTCARDYGAWGCGTTVYSSGGG 120  
QY 121 SGGGSGGGGSDIELTQSPAIMSASPGERVMTCSASSSVRYNMWFQKSGTSPKRWIYD 180  
DB 121 SGGGSGGGGSDIELTQSPAIMSASPGERVMTCSASSSVRYNMWFQKSGTSPKRWIYD 180  
QY 181 TSKLSSGVPARFSSGSGSTSYSLTSSMEADAAATYYCOQWSSNPLTFGAGTKLEKRAA 240  
DB 181 TSKLSSGVPARFSSGSGSTSYSLTSSMEADAAATYYCOQWSSNPLTFGAGTKLEKRAA 240  
QY 241 A 241  
DB 241 A 241  
RESULT 8  
AAR32842  
ID AAR32842 standard; protein; 236 AA.  
XX  
XX AAR32842;  
AC  
XX  
XX 25-MAR-2003 (revised)  
DT 19-JUN-1993 (first entry)

XX  
DE VH NQ10/12.5-VK NQ10/12.5 linked peptide sequences #2.  
XX  
XX Primer; human; immunoglobulin; Ig; variable region; VH; VL; Ck; JH;  
XX lymphocyte; vector; soluble; antibody; phage; linker; back; VH3; nested;  
XX in-cell PCR; cloning; polymorphic; TCR V; antiphenyloxazolone; hybridoma;  
XX NQ2/12.4; NQ10/12.5.  
XX  
XX  
XX OS Synthetic.  
XX  
XX  
XX FH Key Location/Qualifiers  
FT Region 1.115  
FT /label= VH\_NQ10/12.5  
FT Peptide 116..129  
FT /note= "Linker peptide"  
FT Region 130..236  
FT /label= Vkappa\_NQ10/12.5  
XX  
XX PN WO9303151-A1.  
XX  
XX PD 18-FEB-1993.  
XX  
XX PF 10-AUG-1992; 92WO-GB001483.  
XX  
XX PR 10-AUG-1991; 91GB-00017352.  
XX PR 11-JUN-1992; 92GB-00012419.  
XX  
XX PA (MED1-) MEDICAL RES COUNCIL.  
XX  
XX PI Embleton MJ, Gorochev G, Jones PT, Winter GP;  
XX  
XX DR MPI; 1993-076508/09.  
XX DR N-PSDB; AAK37461.  
XX  
XX PT Treatment of cell populations, partic. hybridomas - to link together  
XX PT copies of 2 or more non-contiguous DNA sequences to facilitate analysis.  
XX  
XX PS Disclosure; Fig 4; 72pp; English.  
XX  
XX CC The sequences given in AAR32840-43 show the mature heavy chain VH domains  
XX CC and the VH light chain genes of the antiphenyloxazolone hybridomas  
XX CC NQ2/12.4 and NQ10/12.5 which have been linked via a linker peptide by in-  
XX CC cell PCR. The cDNA encoding these peptides was synthesised using forward  
XX CC primers annealing to the Ck gene and the JH segment, followed by assembly  
XX CC with linker primers, VH back primers based on the VH3 leader sequence and  
XX CC a forward Ck primer nested in respect to the primer used for cDNA. The  
XX CC assembled product within the cells is then amplified with nested primers  
XX CC annealing to the 5' end of the VH gene and the 3' end of the Jk segment.  
XX CC In-cell PCR may be used to determine gene linkage analysis, particularly  
XX CC for the cloning of gene combinations that are polymorphic within a  
XX CC population of cells, such as the rearranged genes for Ig or TCR V  
XX CC regions. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 236 AA;  
Query Match 73.6%; Score 1194; DB 2; Length 236;  
Best Local Similarity 96.2%; Pred. No. 1e-77; 5; Indels 0; Gaps 0;  
Matches 226; Conservative 4; Mismatches 5;  
QY 3 EYVLQESGGGLVOPGSGSRKLSCAASGFTFSSFGMHWRQAPKGLQEMVAIYSSGSSSTIY 62  
DB 1 DQVLVBSGGGLVOPGSGSRKLSCAASGFTFSSFGMHWRQAPKGLQEMVAIYSSGSSSTIY 60  
QY 63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDVTMYTCARDYGAWGCGTTVYSSGGGSG 122  
DB 61 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDVTMYTCARDYGAWGCGTTVYSSGGGSG 120  
QY 123 GGGSGGGGSDIELTQSPAIMSASPGERVMTCSASSSVRYNMWFQKSGTSPKRWIYDTS 182  
DB 121 GGGSGGGGSDIELTQSPAIMSASPGERVMTCSASSSVRYNMWFQKSGTSPKRWIYDTS 180  
QY 183 KLSGVPARFSSGSGSTSYSLTSSMEADAAATYYCOQWSSNPLTFGAGTKLEK 237

|                           |  |        |   |     |
|---------------------------|--|--------|---|-----|
| Dd                        |  | 181    | KLSSGVPARFGSGGCTSYSLTISMSMEDATYYCQOMSSNPILFPGACTKLEK                      | 235 |
|                           |  |        | RESULT 9  |     |
|                           |  |        | AAR68613  |     |
| ID                        |  |        | AAR68613 standard; protein; 240 AA.                                       |     |
| XX                        |  |        |   |     |
| AC                        |  |        | AAR68613;   |     |
| XX                        |  |        |   |     |
| DT                        |  |        | 25-MAR-2003 (revised)   |     |
| DT                        |  |        | 13-SEP-1995 (first entry)   |     |
| XX                        |  |        |   |     |
| DE                        |  |        | Single chain antibody (scfv) which binds to phenylloxazalone.             |     |
| XX                        |  |        |   |     |
| KW                        |  |        | Genetic selection; ligand binding protein; cholera toxin; promoter;       |     |
| KM                        |  |        | detection; selection; beta galactosidase; lac; transmembrane domain;      |     |
| regulatory domain; ds.    |  |        |   |     |
| XX                        |  |        | Synthetic.  |     |
| OS                        |  |        |   |     |
| XX                        |  |        | DEA319296-AI.   |     |
| PN                        |  |        |   |     |
| XX                        |  |        | 15-DEC-1994.  |     |
| PD                        |  |        |   |     |
| PF                        |  |        | 10-JUN-1993; 93DE-04319236.   |     |
| PR                        |  |        | 10-JUN-1993; 93DE-04319236.   |     |
| PA                        |  |        | (BEHM ) BEHRINGWERKE AG.  |     |
| XX                        |  |        |   |     |
| PI                        |  |        | Fritz H, Hennecke F, Kolmar H;  |     |
| DR                        |  |        | WPI, 1995-023689/04.  |     |
| DR                        |  |        | N-PSDB; AA080468.   |     |
| XX                        |  |        |   |     |
| PT                        |  |        | Genetic selection of ligand binding proteins in microorganisms - by       |     |
| PT                        |  |        | extracytoplasmic protein presentation, then use of ligand binding to      |     |
| PT                        |  |        | express a detectable or selectable function.                              |     |
| XX                        |  |        |   |     |
| XX                        |  |        | Example 2.2; Fig 4; 25pp; German.   |     |
| PS                        |  |        |   |     |
| CC                        |  |        | Genetic selection in microorganisms, for ligand binding proteins (LBP)    |     |
| CC                        |  |        | comprises: extracytoplasmic presentation of LBP and; using the signal of  |     |
| CC                        |  |        | ligand binding (by signal transduction on the biosynthetic machinery of   |     |
| CC                        |  |        | the microorganisms) to express a detectable or selectable function.       |     |
| CC                        |  |        | Microorganisms for this process include a genetically stable              |     |
| CC                        |  |        | detection/selection and are transformed with a replicon encoding a fusion |     |
| CC                        |  |        | protein consisting of the LBP, a transmembrane helix and regulatory       |     |
| CC                        |  |        | domain. The detection/selection function is expression of a beta-         |     |
| CC                        |  |        | galactosidase gene, integrated into the chromosome and under the control  |     |
| CC                        |  |        | of the ctx (cholera toxin) promoter. The transmembrane helix is taken     |     |
| CC                        |  |        | from the toxR gene. Four primers (AA080457-60) were used in the           |     |
| CC                        |  |        | construction of the plasmid pHToxscfv. The primers described in AA080457  |     |
| CC                        |  |        | -58 were used to amplify variable heavy chain sequence of the single      |     |
| CC                        |  |        | chain antibody NO10.12.5 and those described in AA080459-60 were used to  |     |
| CC                        |  |        | amplify the corresponding light chain sequence from the same antibody.    |     |
| CC                        |  |        | The amplified sequences were cloned into the plasmid pHToxREI (See        |     |
| CC                        |  |        | AA080454-56 for details) to create a toxR-scFv fusion gene. This sequence |     |
| CC                        |  |        | is the single chain antibody (scfv). (Updated on 25-MAR-2003 to correct   |     |
| CC                        |  |        | PN field.)  |     |
| XX                        |  |        |   |     |
| XX                        |  |        |   |     |
| SQ                        |  |        | Sequence 240 AA;  |     |
|                           |  |        |   |     |
| Query Match               |  | 73.1%; | Score 1185; DB 2; Length 240;   |     |
| Best local Similarity     |  | 95.7%; | Pred. No. 4, 6e-77;   |     |
| Matches 225; Conservative |  | 2;     | Mismatches 8; Indels 0; Gaps 0  |     |
| Oy                        |  |        | 4 VKLDSEGGGLVOPGGSRKLSCAAGFTSSFGMHWRQAPEKGLEMVAYISSGSSITYYA 63            |     |
|                           |  |        | : :   |     |
| Dd                        |  |        | 6 VOLVELGGGFVOPGGSRKLSCAAGFTSSFGMHWRQAPEKGLEMVAYISSGSSITYYA 65            |     |
|                           |  |        | 64 DTVGGRFTISRDNPKNTLLFLQMTSIRSEDVTWYICARDYGAWGCCTTYTVSSGGCGSGC 123       |     |

|           |   |   |     |  |  |
|-----------|---|---|-----|--|--|
| Db        | 66  | DTVKRFRFTSRDNPKNITFLQWTLRSRSEDPTAMYYCARDYGAIWGQTLVTYSSGGGGSGG | 123 |  |  |
| Qy        | 124   | GGSGGGGSDIELTQSPALMSASPGERVTWTCGASSSVRYMNFQOKSGTSPKRWIYDTSK   | 163 |  |  |
| Db        | 126   | GGSGGGGSDIVLTQSPALMSASPGKVTMTTCGASSSVRYMNFQOKSGTSPKRWIYDTSK   | 165 |  |  |
| Qy        | 184   | LSGCVPARFSGSGSGTSTSLTSSNHEADDAATYYCCQMSNPLTPAGTKLEIKR         | 238 |  |  |
| Db        | 186   | LSGCVPARFSGSGSGTSTSLTSSNHEADDAATYYCCQMSNPLTPAGTKLEIKR         | 240 |  |  |
| RESULT 10 |   |   |     |  |  |
| ID        | AAK32841  | AAK32841 standard; protein; 223 AA.                           |     |  |  |
| AC        | AAK32841;   |   |     |  |  |
| XX        |   |   |     |  |  |
| DT        | 25-MAR-2003   | (revised)   |     |  |  |
| DT        | 19-JUN-1993   | (first entry)   |     |  |  |
| XX        |   |   |     |  |  |
| DE        | VH NQ10/12.5-VK NQ10/12.5 linked peptide sequences.                         |   |     |  |  |
| XX        |   |   |     |  |  |
| KW        | Primer: human; immunoglobulin; Ig; variable region; VH; VL; Ck; JH;         |   |     |  |  |
| KW        | lymphocyte; vector; soluble; antibody; phage; linker; back; VH3; nested;    |   |     |  |  |
| KW        | in-cell PCR; cloning; polymorphic; TCR V; antiphenyloxazalone; hybridoma;   |   |     |  |  |
| KW        | NQ2/12.4; NQ10/12.5.  |   |     |  |  |
| XX        |   |   |     |  |  |
| OS        | Synthetic.  |   |     |  |  |
| XX        |   |   |     |  |  |
| PH        | Key   | Location/Qualifiers   |     |  |  |
| FT        | Region  | 1..114  |     |  |  |
| FT        |   | /label= VH_NQ10/12.5  |     |  |  |
| FT        | Peptide   | 115..116  |     |  |  |
| FT        |   | /note= "Linker peptide"                                       |     |  |  |
| FT        | Region  | 117..223  |     |  |  |
| FT        |   | /label= Vkappa_NQ10/12.5                                      |     |  |  |
| XX        |   |   |     |  |  |
| PN        | WO9303151-A1.   |   |     |  |  |
| XX        |   |   |     |  |  |
| PD        | 18-FEB-1993.  |   |     |  |  |
| XX        |   |   |     |  |  |
| PF        | 10-AUG-1992;  | 92MO-GB001483.  |     |  |  |
| XX        |   |   |     |  |  |
| PR        | 10-AUG-1991;  | 91GB-00017352.  |     |  |  |
| PR        | 11-JUN-1992;  | 92GB-00012419.  |     |  |  |
| XX        |   |   |     |  |  |
| PA        | (MEDI-) MEDICAL RES COUNCIL.  |   |     |  |  |
| XX        |   |   |     |  |  |
| PI        | Embleton MJ, Gorochov G, Jones PT, Winter GP;                               |   |     |  |  |
| XX        |   |   |     |  |  |
| DR        | WPI, 1993-076508/09.  |   |     |  |  |
| XX        |   |   |     |  |  |
| XX        | N-PSDB; AAQ37460.   |   |     |  |  |
| PT        | Treatment of cell populations, partic. hybridomas - to link together        |   |     |  |  |
| XX        | copies of 2 or more non-contiguous DNA sequences to facilitate analysis.    |   |     |  |  |
| PS        | Disclosure; Fig 3; 72pp; English.   |   |     |  |  |
| XX        |   |   |     |  |  |
| CC        | The sequences given in AAK32840-43 show the mature heavy chain VH domains   |   |     |  |  |
| CC        | and the VK light chain genes of the antiphenyloxazalone hybridomas          |   |     |  |  |
| CC        | NQ2/12.4 and NQ10/12.5 which have been linked via a linker peptide by in-   |   |     |  |  |
| CC        | cell PCR. The cDNA encoding these peptides was synthesised using forward    |   |     |  |  |
| CC        | primers annealing to the Ck gene and the JH segment, followed by assembly   |   |     |  |  |
| CC        | with linker primers. VH back primers based on the VH3 leader sequence and a |   |     |  |  |
| CC        | forward Ck primer nested in respect to the primer used for cDNA. The        |   |     |  |  |
| CC        | assembled product within the cells is then amplified with nested primers    |   |     |  |  |
| CC        | annealing to the 5' end of the VH gene and the 3' end of the Jk segment.    |   |     |  |  |
| CC        | In-cell PCR may be used to determine gene linkage analysis, particularly    |   |     |  |  |
| CC        | for the cloning of gene combinations that are polymorphic within a          |   |     |  |  |
| CC        | population of cells, such as the rearranged genes for Ig or TCR V           |   |     |  |  |
| CC        | regions. (Updated on 25-MAR-2003 to correct PN field.)                      |   |     |  |  |
| XX        |   |   |     |  |  |
| 90        | Sequence 223 AA;  |   |     |  |  |



KM B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2; CD86;  
KM T cell activation; inhibitor; graft versus host disease;  
KM transplant rejection; allograft rejection; autoimmune disease; allergy;  
KM therapy; human; bispecific tetraivalent antibody; B7Ab;  
KM B7Ab1G10-B7-24H6.  
XX Mus sp.  
OS Homo sapiens.  
OS Synthetic.  
OS Chimeric.  
XX Key  
FH Location/Qualifiers  
FT 1..120  
FT /note= "VH region anti B7.2 MAb"  
FT Peptide  
FT 121..135  
FT /note= "G4S3 flexible linker"  
FT Region  
FT 136..248  
FT /note= "VL region anti B7.2 MAb"  
FT Region  
FT 249..259  
FT /note= "human IgG3 hinge region"  
FT Domain  
FT 260..285  
FT /note= "helix-turn-helix dimerisation domain"  
FT Domain  
FT 286..305  
FT /note= "human IgG3 hinge domain"  
FT Region  
FT 306..426  
FT /note= "VH region anti B7.1 MAb"  
FT Peptide  
FT 427..441  
FT /note= "G4S3 flexible linker"  
FT Region  
FT 442..550  
FT /note= "VL region anti B7.1 MAb"  
FT Peptide  
FT 551..556  
FT /note= "His6 tag"  
XX WO9858965-A2.  
XX 30-DEC-1998.  
XX PD  
XX 22-JUN-1998; 98WO-EP003791.  
XX PR 20-JUN-1997; 97EP-00870092.  
XX PA (INNO-) INNOGENETICS NV.  
XX PI Lorre K, Sablon E, Buyse M, Boeman A;  
XX WPI; 1999-105615/09.  
XX DR N-PSDB; AAX01652.  
XX PT New molecules which bind B7.1 and B7.2 - useful to prevent and treat  
XX immune diseases including allograft rejection.  
XX Example 7.1; Fig 18; 182pp; English.  
XX PS  
XX This polypeptide comprises the bispecific tetraivalent antibody B7Ab1G10-  
CC B7-24H6. The molecule consists of 4 scFvs, i.e. 2 anti B7.1 scFvs and 2  
CC anti B7.2 scFvs (tetraavalency). One single B7Ab is a homodimer of 2  
CC identical molecules, each containing both an anti B7.1 and anti B7.2 scFv  
CC (bispecificity). An anti-B7.1 and anti-B7.2 scFv are linked using a  
CC dimerisation domain (see AAW90219), which drives the homodimerisation of  
CC the molecule. DNA (see AAX01652) encoding the B7Ab has been constructed  
CC to allow expression of the B7Ab in transformed E. coli cells. The B7Ab  
CC crosses links, and/or crosses-reacts, with the costimulatory molecules B7.1  
CC and B7.2 that are expressed on the membrane of professional antigen-  
CC presenting cells, leading to the inhibition of antigen-specific T cell  
CC activation. The invention relates to such B7-binding molecules, methods  
CC for their production, and their use for treating or preventing diseases  
CC of the immune system, in particular graft rejection, graft versus host  
CC disease, allergy and autoimmune diseases (Claimed)  
XX  
XX Sequence 556 AA;  
XX  
XX Query Match 61.7%; Score 1001; DB 2; Length 556;  
XX Best Local Similarity 79.3%; Pred. No. 2e-63;

Matches 192; Conservative 20; Mismatches 26; Indels 4; Gaps 2;  
QY 1 MAEVLQSGGGGLVPGGSRRLSCAASGFTSSFGMHWVROAPEKLEWAVISSGSTI 60  
DB 309 LVQVQLQSGGGLVPGGSRRLSCAASGFTSSFGMHWVROAPEKLEWAVISSSTLI 368  
QY 61 YYADTVKGRFTISRDNPKNLTFLQWTSIRSEDTYVNYCARD--YQAVYGGQGTIVVSSGG 118  
DB 369 YYADSVKGRFTISRDNAKNTLFLOMNSLRADDTAVYCARQGWYPDVWGQGTIVVSSGG 428  
QY 119 GSGGGGSGGGSDIELTQSPAIMSASPGERVMTCSASSSVRYMN--WFOQSGTSPKR 176  
DB 429 GSGGGGSGGGSDIELTQSPSMAASVGDRTITCSYSSRLSSSLHMYQKSTSPRP 488  
QY 177 WIYDTSKLSSGVPAFSSGSGTSTSLTSSMEADAAATYYCOQSSNPLTFAGTKLEL 236  
DB 489 WIYGTSLNASGVPSRFSGSGGTDTLTSSMQPEDAATYYCOQSSYPLTFGQGTLEI 548  
QY 237 KR 238  
DB 549 KR 550  
RESULT 13  
AAE38657  
ID AAE38657 standard; protein, 237 AA.  
XX  
XX AAE38657;  
AC  
XX  
XX 04-DEC-2003 (first entry)  
DT  
XX  
XX Mouse G1 single chain Fv-recombinant antibody.  
DE  
XX  
XX Major histocompatibility complex; MHC; HLA-restricted antigen; cancer;  
KM viral infection; autoimmune disease; gene therapy; cytostatic; virucide;  
KM immunomodulator; mouse.  
XX  
XX Mus sp.  
OS  
XX  
XX Key  
FH Location/Qualifiers  
FT 1..118  
FT Domain  
FT /note= "VH domain"  
FT 31..35  
FT /note= "Complementarity determining region"  
FT 50..66  
FT /note= "Complementarity determining region"  
FT 99..108  
FT /note= "Complementarity determining region"  
FT 119..234  
FT /note= "Peptide linker"  
FT Domain  
FT 135..237  
FT /note= "VL domain"  
FT 158..167  
FT /note= "Complementarity determining region"  
FT 182..189  
FT /note= "Complementarity determining region"  
FT 222..230  
FT /note= "Complementarity determining region"  
XX WO2003068201-A2.  
XX 21-AUG-2003.  
XX PD  
XX 11-FEB-2003; 2003WO-IL000105.  
XX PF  
XX 13-FEB-2002; 2002US-00073301.  
XX PR  
XX (TECR ) TECHNION RES & DEV FOUND LTD.  
XX PA  
XX Reiter Y, Denkerberg G;  
XX  
XX WPI; 2003-689603/65.  
XX DR N-PSDB; AAD58607.

XX New isolated molecule comprising an antibody that binds with a human  
 PT major histocompatibility complex (MHC) class I being complexed with a HLA  
 PT -restricted antigen, useful for treating cancer, viral infection or  
 PT autoimmune disease.  
 XX  
 PS Claim 63; Fig 3a; 81pp; English.  
 XX  
 CC The invention relates to an isolated molecule comprising an antibody  
 CC specifically bindable with a binding affinity below 20 nanomolar to a  
 CC human major histocompatibility complex (MHC) class I being complexed with  
 CC a HLA-restricted antigen. The molecules, antibodies, and methods are  
 CC useful for treating cancer, viral infection and an autoimmune disease.  
 CC The invention is useful in gene therapy. The present sequence is mouse G1  
 CC single chain Fv-recombinant antibody  
 XX  
 SO Sequence 237 AA;  
 XX  
 Query Match 60.2%; Score 976.5; DB 7; Length 237;  
 Best Local Similarity 78.9%; Pred. No. 4.1e-62;  
 Matches 187; Conservative 18; Mismatches 27; Indels 5; Gaps 1;  
 QY 3 EVKLQSGGGGLVOPGSSRKLSCAASGFTFSFGMHWROAPEKLEWVAIYSSGSTIYY 62  
 DB 1 QVQLQSGGGGLVPGGSLKLSCAASGFTFSFGMHWROAPEKLEWVAIYSSGSTIYY 60  
 QY 63 ADTVKGRFTISRDPKNTLFLQWTSLSRSDPTVWYVCARD-----YGAHYGQGTYYVSSG 117  
 DB 61 PDSVKGRFTISRDPKNTLFLQWTSLSRSDPTVWYVCARDGMEGWYDVGQGTYYVSSG 120  
 QY 118 GGGSGGGGGGGGSDIELTQSPAIMSASPERVTMTCSASSSVRYNMWFQOKSGTSPKRW 177  
 DB 121 GGGSGGGGGGGGSDIELTQSPAIMSASPERVTMTCSASSSIRIYVWQKXGSSPRL 180  
 QY 178 IYDTSKLSSGVPARFSGSGSGTSLTISMEADDAIYYCQWSSNPLTFGAGT 234  
 DB 181 IYDTSNVAGVPRFSGSGSGTSLTINMEADDAIYYCQWSSNPLTFGAGT 237

RESULT 14  
 AAY44973  
 ID AAY44973 standard; protein; 268 AA.  
 AC AAY44973;  
 XX  
 XX 23-MAY-2000 (first entry)  
 DT  
 XX  
 DE Recombinant mouse anti-rotavirus antibody (Clone 22).  
 XX  
 XX Recombinant anti-rotavirus antibody; mouse; heavy chain variable region;  
 KM VH; light chain variable region; VL; PCANTAB 5E vector; treatment;  
 KM probiotic bacteria; antibacterial; antiviral; gastrointestinal tract;  
 KM immunise; neonate; immunosuppressed; immunodeficient; GI pathogen.  
 XX  
 XX Synthetic.  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT 6..116  
 FT /label= Mouse\_Heavy\_chain\_variable\_region  
 FT 117..152  
 FT /label= Linker\_region  
 FT 153..250  
 FT /label= Mouse\_Light\_chain\_variable\_region  
 FT 254..266  
 FT /label= B\_tag\_domain  
 FT /note= "Snables immunodetection and immunosensitivity  
 FT purification of the recombinant antibody"  
 XX  
 XX MO200006764-A1.  
 XX  
 XX 10-FEB-2000.  
 XX

PF 29-JUL-1999; 99WO-US017296.  
 XX  
 XX 30-JUL-1998; 98US-0094697P.  
 XX  
 PA (WISC) WISCONSIN ALUMNI RES FOUND.  
 XX  
 PI Fahl WE, Letchworth GJ, Mueller GC, Savage AK, Loo D;  
 DR WPI; 2000-195315/17.  
 DR N-PSDB; AA250509.  
 XX  
 PT Composition for supplementing or replacing an immune response against  
 PT gastrointestinal pathogens in e.g. newborn infants, comprises probiotic  
 PT microorganisms expressing antibodies specific for the gastrointestinal  
 PT pathogens.  
 XX  
 XX Disclosure; Fig 2; 48pp; English.  
 XX  
 CC The present sequence is the recombinant mouse anti-rotavirus antibody. It  
 CC comprises of mouse heavy chain (VH) and light chain (VL) variable regions  
 CC joined by a linker. The recombinant antibody encoding DNA is inserted  
 CC into PCANTAB 5E expression vector (clone 22). The vector is used to  
 CC transform probiotic bacteria like, Lactobacilli for expression and  
 CC secretion of recombinant anti-rotavirus antibodies. The antibodies have  
 CC antibacterial and antiviral activity. Antibody-expressing probiotic  
 CC bacteria can be administered for treatment of gastrointestinal (GI) tract  
 CC infections and to immunise neonates, humans or immunosuppressed/  
 CC immunodeficient adults acutely exposed to a bolus of GI pathogen  
 XX  
 SO Sequence 268 AA;  
 XX  
 Query Match 59.8%; Score 970.5; DB 3; Length 268;  
 Best Local Similarity 74.7%; Pred. No. 1.3e-61;  
 Matches 186; Conservative 24; Mismatches 30; Indels 9; Gaps 3;  
 QY 1 MAEVLQSGGGGLVOPGSSRKLSCAASGFTFSFGMHWROAPEKLEWVAIYSSGSTI 60  
 DB 6 MAQVQLQSGGGGLVPGGSLKLSCAASGFTFSFGMHWROAPEKLEWVAIYSSGSTI 65  
 QY 61 YPADTVKGRFTISRDPKNTLFLQWTSLSRSDPTVWYVCARDG-----AYWGQGTYYT 113  
 DB 66 HYADTVKGRFTISRDPKNTLFLQWTSLSRSDPTVWYVCARDGMEGWYDVGQGTYYT 124  
 QY 114 VSSGGGGGGGGGSDIELTQSPAIMSASPERVTMTCSASSSVRYNMWFQOKSGT 172  
 DB 125 VSSGGGGGGGGGSDIELTQSPAIMSASPERVTMTCSASSSVRYNMWFQOKSGT 184  
 QY 173 SPKRWIYDTSKLSGVPARFSGSGSGTSLTISMEADDAIYYCQWSSNPLTFGAGT 232  
 DB 185 SPQFLVYSAKTLAEGVPSRFSGSGSGTQSLKINSIQPDPFGNYCQHYGSPRTFGAGT 244  
 QY 233 KLEIKRAAA 241  
 DB 245 KLEIKRAAA 253

RESULT 15  
 AAY44972  
 ID AAY44972 standard; protein; 268 AA.  
 AC AAY44972;  
 XX  
 XX 23-MAY-2000 (first entry)  
 DT  
 XX  
 DE Recombinant mouse anti-rotavirus antibody (Clone 11).  
 XX  
 XX Recombinant anti-rotavirus antibody; mouse; heavy chain variable region;  
 KM VH; light chain variable region; VL; PCANTAB 5E vector; treatment;  
 KM probiotic bacteria; antibacterial; antiviral; gastrointestinal tract;  
 KM immunise; neonate; immunosuppressed; immunodeficient; GI pathogen.  
 XX  
 XX Synthetic.  
 OS Mus sp.  
 OS



Search completed: June 23, 2005, 05:53:56  
Job time : 82 secs

```
XX Key Location/Qualifiers
FH Region 6..116
FT Region /label= Mouse_Heavy_chain_variable_region
FT Region 117..151
FT Region /label= Linker_region
FT Region 152..250
FT Domain /label= Mouse_Light_chain_variable_region
FT Domain 254..266
FT /label= B tag domain
FT /note= "Enables immunodetection and immunoaffinity
FT purification of the recombinant antibody"
XX
XX WO200006764-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US017296.
XX
XX 30-JUL-1998; 98US-0094697P.
XX
XX (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
XX Fahl WE, Letchworth GJ, Mueller GC, Savage AK, Loo D;
XX
XX WPI; 2000-195315/17.
XX
XX N-PSDB; AA250508.
XX
XX Composition for supplementing or replacing an immune response against
XX gastrointestinal pathogens in e.g. newborn infants, comprises probiotic
XX microorganisms expressing antibodies specific for the gastrointestinal
XX pathogens.
XX
XX Disclosure; Fig 1; 48p; English.
XX
XX
XX The present sequence is the recombinant mouse anti-rotavirus antibody. It
XX comprises of mouse heavy chain (VH) and light chain (VL) variable regions
XX joined by a linker. The recombinant antibody encoding DNA is inserted
XX into pCMV7AB 5E expression vector (clone 11). The vector is used to
XX transform probiotic bacteria like, Lactobacilli for expression and
XX secretion of recombinant anti-rotavirus antibodies. The antibodies have
XX antibacterial and antiviral activity. Antibody-expressing probiotic
XX bacteria can be administered for treatment of gastrointestinal (GI) tract
XX infections and to immunise neonates, humans or immunosuppressed/
XX immunodeficient adults acutely exposed to a bolus of GI pathogen
XX
XX
XX Sequence 268 AA;
SQ
Query Match 59.6%; Score 967.5; DB 3; Length 268;
Best Local Similarity 75.1%; Pred. No. 2.1e-61;
Matches 187; Conservative 22; Mismatches 31; Indels 9; Gaps 3;
QY 1 MAEVLQESGGGLVOPGSGRKLSCAAGFTFSPFGHWYRQAPKGLWVAYISSGSTI 60
DB ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 60
QY 6 MAQVQLQDSGGGLVOPGSGRKLSCAAGFTFSPFGHWYRQAPKGLWVAYISSGSTI 65
DB ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 65
QY 61 VYADTVKGRFTISRDNPKNLFLQMTSLRSEDTVMYTCARDYG-----AYMGQTIYT 113
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 113
QY 66 HYADTVKGRFTISRDNPKNLFLQMTSLRSEDTVMYTCARDYG-----AYMGQTIYT 124
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 124
QY 114 VSSGGGGSGGGGGSGGSDIELTQSPALMSAPGERVTMTCSASSSV-RYNNWFOQKSGT 172
DB ||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 172
QY 125 VSSGGGGSGGGGGSGGSDIVLTQSPALMSAPGERVTMTCSASSSV-RYNNWFOQKSGT 184
DB ||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 184
QY 173 SPKRWIYDTSKLSGVPARFSGSGSTSYSLTISWEADATYYCQQWSNPLTFGAGT 232
DB ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 232
QY 185 SPQFLVTSKXTLAEGVPSRFSGSGSTQPSLKINSIQPDPFGNYICQHYGTPRTFGAGT 244
DB ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 244
QY 233 KLELKRAAA 241
DB |||||||
QY 245 KLELKRAAA 253
DB |||||||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 05:52:38 ; Search time 1196 Seconds  
(without alignments)  
101.603 Million cell updates/sec

Title: US-09-403-882a-2

Perfect score: 1622

Sequence: 1 MAEVLQESGGGLVQPGGSR.....VLLTITLILIMLQKKR 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US10F\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID                 | Description        |
|------------|-------|-------------|--------|--------------------|--------------------|
| 1          | 959.5 | 59.1        | 291    | US-10-406-830-10   | Sequence 10, Appl  |
| 2          | 943.5 | 58.2        | 288    | US-09-818-247-22   | Sequence 22, Appl  |
| 3          | 940   | 58.0        | 244    | US-10-879-994-16   | Sequence 16, Appl  |
| 4          | 928   | 57.2        | 244    | US-10-610-452-16   | Sequence 16, Appl  |
| 5          | 927   | 57.2        | 291    | US-10-805-177-111  | Sequence 111, Appl |
| 6          | 926   | 57.1        | 287    | US-10-406-830-9    | Sequence 9, Appl   |
| 7          | 913   | 56.3        | 282    | US-10-406-830-5    | Sequence 5, Appl   |
| 8          | 913   | 56.3        | 242    | US-10-259-087A-20  | Sequence 20, Appl  |
| 9          | 907   | 55.9        | 242    | US-10-689-006-20   | Sequence 20, Appl  |
| 10         | 906.5 | 55.9        | 353    | US-09-203-958A-4   | Sequence 4, Appl   |
| 11         | 906.5 | 55.9        | 237    | US-09-880-748-2104 | Sequence 2104, Ap  |

|    |       |      |     |    |                    |                   |
|----|-------|------|-----|----|--------------------|-------------------|
| 12 | 906.5 | 55.9 | 237 | 15 | US-10-293-418-2104 | Sequence 2104, Ap |
| 13 | 903.5 | 55.7 | 237 | 10 | US-09-880-748-2020 | Sequence 2020, Ap |
| 14 | 903.5 | 55.7 | 237 | 15 | US-10-293-418-2020 | Sequence 2020, Ap |
| 15 | 900.5 | 55.5 | 237 | 10 | US-09-880-748-2005 | Sequence 2005, Ap |
| 16 | 900.5 | 55.5 | 237 | 10 | US-09-880-748-2114 | Sequence 2114, Ap |
| 17 | 900.5 | 55.5 | 237 | 15 | US-10-293-418-2005 | Sequence 2005, Ap |
| 18 | 900.5 | 55.5 | 237 | 15 | US-10-293-418-2114 | Sequence 2114, Ap |
| 19 | 899.5 | 55.5 | 237 | 10 | US-09-880-748-2118 | Sequence 2118, Ap |
| 20 | 899.5 | 55.5 | 237 | 15 | US-10-293-418-2118 | Sequence 2118, Ap |
| 21 | 899.5 | 55.5 | 639 | 17 | US-10-792-498-16   | Sequence 16, Appl |
| 22 | 898.5 | 55.4 | 237 | 10 | US-09-880-748-2115 | Sequence 2115, Ap |
| 23 | 898.5 | 55.4 | 237 | 15 | US-10-293-418-2115 | Sequence 2115, Ap |
| 24 | 897.5 | 55.3 | 237 | 10 | US-09-880-748-2027 | Sequence 2027, Ap |
| 25 | 897.5 | 55.3 | 237 | 10 | US-09-880-748-2111 | Sequence 2111, Ap |
| 26 | 897.5 | 55.3 | 237 | 15 | US-10-293-418-2027 | Sequence 2027, Ap |
| 27 | 897.5 | 55.3 | 237 | 15 | US-10-293-418-2111 | Sequence 2111, Ap |
| 28 | 897.5 | 55.3 | 203 | 15 | US-10-239-656-77   | Sequence 77, Appl |
| 29 | 897   | 55.3 | 256 | 15 | US-10-239-656-61   | Sequence 61, Appl |
| 30 | 896.5 | 55.3 | 237 | 10 | US-09-880-748-2003 | Sequence 2003, Ap |
| 31 | 896.5 | 55.3 | 237 | 10 | US-09-880-748-2017 | Sequence 2017, Ap |
| 32 | 896.5 | 55.3 | 237 | 10 | US-09-880-748-2019 | Sequence 2019, Ap |
| 33 | 896.5 | 55.3 | 237 | 10 | US-09-880-748-2028 | Sequence 2028, Ap |
| 34 | 896.5 | 55.3 | 237 | 10 | US-09-880-748-2110 | Sequence 2110, Ap |
| 35 | 896.5 | 55.3 | 237 | 15 | US-10-293-418-2003 | Sequence 2003, Ap |
| 36 | 896.5 | 55.3 | 237 | 15 | US-10-293-418-2017 | Sequence 2017, Ap |
| 37 | 896.5 | 55.3 | 237 | 15 | US-10-293-418-2019 | Sequence 2019, Ap |
| 38 | 896.5 | 55.3 | 237 | 15 | US-10-293-418-2028 | Sequence 2028, Ap |
| 39 | 896.5 | 55.3 | 237 | 15 | US-10-293-418-2110 | Sequence 2110, Ap |
| 40 | 895.5 | 55.2 | 237 | 15 | US-09-880-748-1906 | Sequence 1906, Ap |
| 41 | 895.5 | 55.2 | 237 | 15 | US-10-293-418-1906 | Sequence 1906, Ap |
| 42 | 894.5 | 55.1 | 237 | 10 | US-09-880-748-2040 | Sequence 2040, Ap |
| 43 | 894.5 | 55.1 | 237 | 15 | US-10-293-418-2040 | Sequence 2040, Ap |
| 44 | 894   | 55.1 | 240 | 9  | US-09-192-854-2    | Sequence 2, Appl  |
| 45 | 894   | 55.1 | 240 | 9  | US-09-968-561A-2   | Sequence 2, Appl  |

## ALIGNMENTS

RESULT 1  
US-10-406-830-10  
; Sequence 10, Application US/10406830  
; Publication No. US20040071696A1  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, GREGORY P.  
; APPLICANT: HORAK, EVA M.  
; APPLICANT: WEINER, LOUIS M.  
; TITLE OF INVENTION: BISPECIFIC SINGLE CHAIN Fv ANTIBODY MOLECULES AND METHODS OF USE  
; FILE REFERENCE: 407T-000410US  
; CURRENT APPLICATION NUMBER: US/10/406, 830  
; CURRENT FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: US 60/370, 276  
; PRIOR FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 10  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic antibody.  
US-10-406-830-10

Query Match 59.1%; Score 959; DB 15; Length 291;  
Best Local Similarity 70.1%; Pred. No. 2.2e-57;  
Matches 185; Conservative 31; Mismatches 38; Indels 10; Gaps 2;

QY 1 MAEVLQESGGGLVQPGGSRKSLSCASGTFPSFGHWRQAREKGLWAVYSSSSSTI 60  
DB 21 MAOVLOESGGGLVQPGGSRKSLSCASGTFPSFYAMWROAPGKGLWVSAISSGGST 80

```

QY 61 YYADTVKGRFTTISRDNPKNTLFLQMTSLRSEDVTMYTCARD-----YGAWVGQGT 111
DB 81 YYADSVKGRFTTISRDNPKNTLYLQNMNSLRAEDTALYYCARBEGYSNMNMWYFDLMGRGTL 140
QY 112 YVVSOGGSGGGGGGGGGSDIELTOSPAIMASPERVTMTCSASSV-RYMMNPQOQS 170
DB 141 YVVSOGGSGGGGGGGGGSEIVLTOSPSLSASVDRITTCRASQSSISLMTQQXP 200
QY 171 GTSPPKMIYDTSKLSGVPARFSGSGSTYSLTSSMEADPAATYYCOOWSNPLTFGA 230
DB 201 GNAPKLLIYVASSLQSGVPSRFSGSGSGTDFLTITISLQDPDPATYYCQOYNSYPTFGQ 260
QY 231 GTKLEIKRAAAEOKLISEEDLNGA 254
DB 261 GTKLEIKRAAAEOKLISEEDLNGA 284

```

# RESULT 2

```

US-09-818-247-22
/ Sequence 22, Application US/09818247
/ Patent No. US20020102657A1
/ GENERAL INFORMATION:
/ APPLICANT: Mostov, Keith E.
/ APPLICANT: Chapin, Steven J.
/ APPLICANT: Richman-Bienbat, Janice
/ TITLE OF INVENTION: The Regents of the University of California
/ TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Steroid Region of p18r and Methods of Use Ther
/ FILE REFERENCE: 18062E-0009100S
/ CURRENT APPLICATION NUMBER: US/09/818,247
/ PRIOR FILING DATE: 2001-03-26
/ PRIOR APPLICATION NUMBER: WO PCT/US01/09699
/ PRIOR FILING DATE: 2001-03-26
/ PRIOR APPLICATION NUMBER: US 60/192,197
/ PRIOR FILING DATE: 2000-03-27
/ PRIOR APPLICATION NUMBER: US 60/192,198
/ PRIOR FILING DATE: 2000-03-27
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 22
/ LENGTH: 288
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial
/ OTHER INFORMATION: Sequence: pelb/4AF/myc/SHIS
US-09-818-247-22

```

Query Match 58.2%; Score 943.5; DB 9; Length 288;

Best Local Similarity 69.3%; Pred. No. 2.4e-56;

Matches 181; Conservative 32; Mismatches 41; Indels 7; Gaps 2;

```

QY 1 MAEVKLOESGGGLVOPGSRKLSGASGFTFSGFMHWVQAPKGLKLEWVAIYSSGSSSTI 60
DB 21 MAOVQVQSGGGLVOPGSLRLSCASGFTFSSYAMSWRQAPKGLKLEWVAIYSSGSGST 80
QY 61 YYADTVKGRFTTISRDNPKNTLFLQMTSLRSEDVTMYTCARD-----YGAWVGQGT 114
DB 81 YYADSVKGRFTTISRDNPKNTLYLQNMNSLRAEDTALYYCARBEGYSNMNMWYFDLMGR 140
QY 112 YVVSOGGSGGGGGGGGGSDIELTOSPAIMASPERVTMTCSASSV-RYMMNPQOQS 173
DB 141 YVVSOGGSGGGGGGGGGSEIVLTOSPSLSASVDRITTCRASQSSISLMTQQXP 200
QY 171 GTSPPKMIYDTSKLSGVPARFSGSGSTYSLTSSMEADPAATYYCOOWSNPLTFGA 233
DB 201 GNAPKLLIYVASSLQSGVPSRFSGSGSGTDFLTITISLQDPDPATYYCQOYNSYPTFG 260
QY 231 GTKLEIKRAAAEOKLISEEDLNGA 254
DB 261 GTKLEIKRAAAEOKLISEEDLNGA 284

```

```

RESULT 3
US-10-879-994-16
/ Sequence 16, Application US/10879994
/ Publication No. US20050032175A1
/ GENERAL INFORMATION:
/ APPLICANT: Stahl, Neil
/ APPLICANT: Yancopoulos, George D.
/ APPLICANT: Karrow, Margaret
/ APPLICANT: Smith, Eric
/ TITLE OF INVENTION: HIGH AFFINITY FUSION PROTEINS AND THERAPEUTIC AND DIAGNOSTIC METH
/ TITLE OF INVENTION: USE
/ FILE REFERENCE: REG 203E2
/ CURRENT APPLICATION NUMBER: US/10/879,994
/ PRIOR FILING DATE: 2004-06-29
/ PRIOR APPLICATION NUMBER: 10/610,452
/ PRIOR FILING DATE: 2003-06-30
/ NUMBER OF SEQ ID NOS: 112
/ SOFTWARE: FaetsEQ for Windows Version 4.0
/ SEQ ID NO 16
/ LENGTH: 244
/ TYPE: PRT
/ ORGANISM: mus musculus
US-10-879-994-16

```

Query Match 58.0%; Score 940; DB 17; Length 244;

Best Local Similarity 74.6%; Pred. No. 3.6e-56;

Matches 182; Conservative 22; Mismatches 32; Indels 8; Gaps 3;

```

QY 3 EYKLOESGGGLVOPGSRKLSGASGFTFSGFMHWVQAPKGLKLEWVAIYSSGSSSTIY 62
DB 1 DQVLVESGGGLVOPGSRKLSGASGFTFSGFMHWVQAPKGLKLEWVAIYSSGSSSTIY 60
QY 63 ADTVKGRFTTISRDNPKNTLFLQMTSLRSEDVTMYTCARDYGA-----YMGQGTIVYSS 116
DB 61 ADTVKGRFTTISRDNPKNTLFLQMTSLRSEDVTMYTCASRTARABAMYWGQGTIVYSS 120
QY 117 -GGGGSGGGGGGGGGSDIELTOSPAIMASPERVTMTCSASSV-RYMMNPQOXSSTP 174
DB 121 AGGGSGGGGGGGGGSDIELTOSPASLSVGETVITTCRASGNHNTLAWYQOXSSTP 180
QY 175 KRWIDTSKLSGVPARFSGSGSTYSLTSSMEADPAATYYCOOWSNPLTFGAGTK 234
DB 181 QLVVNTATLADGVPSRFSGSGSGTQYSLKINSLOPEDFGSIYQFMSIPTFGGCTY 240
QY 235 ELKR 238
DB 241 ENKR 244

```

# RESULT 4

US-10-610-452-16

/ Sequence 16, Application US/10610452

/ Publication No. US20050074855A1

/ GENERAL INFORMATION:

/ APPLICANT: George D. Yancopoulos

/ APPLICANT: Neil Stahl

/ TITLE OF INVENTION: Cytokine-Specific Fusion Proteins and

/ TITLE OF INVENTION: Therapeutic and Diagnostic Methods For Use

/ FILE REFERENCE: REG 203E

/ CURRENT APPLICATION NUMBER: US/10/610,452

/ PRIOR FILING DATE: 2003-06-30

/ PRIOR APPLICATION NUMBER: 10/787,835

/ PRIOR FILING DATE: 2001-03-22

/ PRIOR APPLICATION NUMBER: PCT/US9/22045

/ PRIOR FILING DATE: 1999-04-19

/ PRIOR APPLICATION NUMBER: 09/313,942

/ PRIOR FILING DATE: 1999-04-19

/ PRIOR APPLICATION NUMBER: 60/101,858

/ PRIOR FILING DATE: 1999-09-25

/ NUMBER OF SEQ ID NOS: 17

/ SOFTWARE: FaetsEQ for Windows Version 4.0

/ SEQ ID NO 16

/ LENGTH: 244



Publication No. US20040071696A1  
GENERAL INFORMATION:  
APPLICANT: ADAMS, GREGORY P.  
APPLICANT: HORAK, EVA M.  
APPLICANT: WEINER, LOUIS M.  
APPLICANT: JAMES, MARKS D.  
TITLE OF INVENTION: BISPECIFIC SINGLE CHAIN Fv ANTIBODY MOLECULES AND METHODS OF USE  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 407T-000410US  
CURRENT APPLICATION NUMBER: US/10/406,830  
CURRENT FILING DATE: 2003-04-04  
PRIOR APPLICATION NUMBER: US 60/370,276  
PRIOR FILING DATE: 2002-04-05  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 5  
LENGTH: 287  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Synthetic antibody.  
US-10-406-830-5

Query Match 57.1%, Score 926; DB 15; Length 287;  
Best Local Similarity 69.6%; Pred. No. 3.8e-55;  
Matches 181; Conservative 32; Mismatches 41; Indels 6; Gaps 3;

QY 1 MAEVKLOESSGGGLVOPGSGSRKLSCAASGFTFSSFGMHVWROAPEKGLWAVAYISSGSSSTI 60  
DB 21 MAQVQLVDSGGSGLVOPGSRRLRSLSCAASGFTFSDYIHWRQAPGKLEMAVAYISDGNKK 80  
QY 61 YYADTVKGRFTISRDPNPKNTLFLQWTSLSRSDTVMYCAR--DYGA--YMGQGTIVYSS 115  
DB 81 YYAAAVKODFTISRDPNPKNTLFLQWTSLSRSDTVMYCAR--DYGA--YMGQGTIVYSS 140  
QY 116 SGGGSGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSV-RYNNMFQOKSGTSPR 174  
DB 141 SGGGSGGGGSGGGSDIQMTQSPSTLSASLSDRVITTRASQSIGSWLAWYQOKRGPAP 200  
QY 175 KMWYDTSKLSGVPARFSGSGSTSYSLTSSMEADATYYCOQWSSNPLTFGAGTKLEL 234  
DB 201 KLLITKASLTLESGLVPRFSGSGSTFTLTISGLQEDPATYYCQKLSYPLTFGGGTIV 260  
QY 235 ELKRAAAEQKLISEEDLNCA 254  
DB 261 ELKRAAAEQKLISEEDLNCA 280

## RESULT 8

US-10-259-087A-20  
Sequence 20, Application US/10259087A  
Publication No. US20030130190A1  
GENERAL INFORMATION:  
APPLICANT: Vanderbilt University  
APPLICANT: Hallahan, Dennis E  
APPLICANT: Qu, Shihua  
TITLE OF INVENTION: IN VIVO PANNING FOR LIGANDS TO RADIATION-INDUCED MOLECULES  
FILE REFERENCE: 1242/47/2  
CURRENT APPLICATION NUMBER: US/10/259,087A  
CURRENT FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: US 60/328123  
PRIOR FILING DATE: 2001-10-03  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20  
LENGTH: 242  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificial antibody ligand number 2  
US-10-259-087A-20

Query Match 56.3%, Score 913; DB 14; Length 242;

Best Local Similarity 71.1%; Pred. No. 2.4e-54;  
Matches 172; Conservative 30; Mismatches 36; Indels 4; Gaps 2;  
QY 1 MAEVKLOESSGGGLVOPGSGSRKLSCAASGFTFSSFGMHVWROAPEKGLWAVAYISSGSSSTI 60  
DB 1 MAQVQLVDSGGSGLVOPGSRRLRSLSCAASGFTFSDYIHWRQAPGKLEMAVAYISDGNKK 80  
QY 61 YYADTVKGRFTISRDPNPKNTLFLQWTSLSRSDTVMYCAR--DYGA--YMGQGTIVYSS 116  
DB 61 KNEFKKKAALTSIDKSSSTAYMEISLTSEDSAVYYCARFGNYGALYMGQGTIVYSS 120  
QY 117 GGGGSGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSVRYNNMFQOKSGTSPR 176  
DB 121 GGGGSGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSVRYNNMFQOKSGTSPR 180  
QY 177 WYDTSKLSGVPARFSGSGSTSYSLTSSMEADATYYCOQWSSNPLTFGAGTKLEL 236  
DB 181 WYGTSLNLSAGVPRFSGSGSTSYSLTSSMEADATYYCOQWSSYPLTFGGGTIVYSS 240  
QY 237 KR 238  
DB 241 KR 242

## RESULT 9

US-10-689-006-20  
Sequence 20, Application US/10689006  
Publication No. US20040191249A1  
GENERAL INFORMATION:  
APPLICANT: Vanderbilt University  
APPLICANT: Hallahan, Dennis E  
APPLICANT: Wernau, Raymond  
TITLE OF INVENTION: PHAGE ANTIBODIES TO RADIATION-INDUCIBLE NEOANTIGENS  
FILE REFERENCE: 1242/72  
CURRENT APPLICATION NUMBER: US/10/689,006  
CURRENT FILING DATE: 2003-10-20  
PRIOR APPLICATION NUMBER: US 09/914,605  
PRIOR FILING DATE: 2001-08-30  
PRIOR APPLICATION NUMBER: US 10/259,087  
PRIOR FILING DATE: 2002-09-27  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 20  
LENGTH: 242  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Artificial antibody ligand number 2  
US-10-689-006-20

Query Match 56.3%, Score 913; DB 16; Length 242;  
Best Local Similarity 71.1%; Pred. No. 2.4e-54;  
Matches 172; Conservative 30; Mismatches 36; Indels 4; Gaps 2;

QY 1 MAEVKLOESSGGGLVOPGSGSRKLSCAASGFTFSSFGMHVWROAPEKGLWAVAYISSGSSSTI 60  
DB 1 MAQVQLVDSGGSGLVOPGSRRLRSLSCAASGFTFSDYIHWRQAPGKLEMAVAYISDGNKK 80  
QY 61 YYADTVKGRFTISRDPNPKNTLFLQWTSLSRSDTVMYCAR--DYGA--YMGQGTIVYSS 116  
DB 61 KNEFKKKAALTSIDKSSSTAYMEISLTSEDSAVYYCARFGNYGALYMGQGTIVYSS 120  
QY 117 GGGGSGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSVRYNNMFQOKSGTSPR 176  
DB 121 GGGGSGGGGSGGGSDIELTQSPAIMSASPERVMTCSASSVRYNNMFQOKSGTSPR 180  
QY 177 WYDTSKLSGVPARFSGSGSTSYSLTSSMEADATYYCOQWSSNPLTFGAGTKLEL 236  
DB 181 WYGTSLNLSAGVPRFSGSGSTSYSLTSSMEADATYYCOQWSSYPLTFGGGTIVYSS 240  
QY 237 KR 238  
DB 241 KR 242



|    |  |     |  |     |
|----|--|-----|--|-----|
| OY |  | 63  | ADTVKGRFTTIRSDNKNLTFLQMTSLRSEDVWYYCARDYGAVWGCGTTVTYSSGGGSGSg     | 122 |
| Dd |  | 61  | ADVXKGRFTTIRDNKNLSLYLQMSLRRAEDTAIVYICARPTTDYWGCGTTVTYSSGGGSGSg   | 120 |
| OY |  | 123 | GGSGGGGGSDIELTOSPAIMASAPGBERVMTTCASSSV-RYNMWFQOKSGTSPKRWIYDT     | 181 |
| Dd |  | 121 | GGSGGGGGSDIYMTCPSSTLSASVDGDRVITTCRASGGISSMILAWYOQKRGAPKVLIYKA    | 180 |
| OY |  | 182 | SKLSSGVAFRRFSGSGSGTSYSTLTISMEAEADAATYYCOOWSNPLTFEGAGTKLEIKR      | 238 |
| Dd |  | 181 | STLESQVSRFRFSGSGSGTGDTFLTLLTSLQPEDEPATIYYCOOYSYSTPWTFFGGGTYLEIKR | 237 |

RESULT 13  
US-09-880-748-2020

```
Sequence 2020, Application US/09880748
Publication No. US20030059937A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunosepecifically Bind BlyS
FILE REFERENCE: PFS23
CURRENT APPLICATION NUMBER: US/09/880,748
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/212,210
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/240,816
PRIOR FILING DATE: 2000-10-17
PRIOR APPLICATION NUMBER: 60/276,248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,379
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/293,499
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 3239
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2020
LENGTH: 237
TYPE: PRF
ORGANISM: Homo sapiens
US-09-880-748-2020
```

|  |         |                       |  |                      |
|--|---------|-----------------------|--|----------------------|
|  |         | Query Match           | 55.7%; Score 903.5; DB 10; Length 237; |                      |
|  |         | Best Local Similarity | 72.6%; Pred. No. 1e-53;                |                      |
|  | Matches | 172; Conservative     | 27; Mismatches                         | 37; Indels 1; Gaps 1 |

RESULT 14  
US-10-293-418-2020

```
Sequence 2020, Application US/10293418
Publication No. US20030223996A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Antibodies that Immunosepecifically Bind BlyS
FILE REFERENCE: PFS23p2
CURRENT APPLICATION NUMBER: US/10/293,418
CURRENT FILING DATE: 2002-11-27
```

|    |  |     |  |     |
|----|--|-----|--|-----|
| OY |  | 3   | EVKLOESGGGVOPCGSRKRLSCAASGFPTSSRCGMHVRAPBKEGLMWAYISSGSTIYY       | 62  |
| Dd |  | 1   | EVVLQVSOGGLVQPFGQSURLSCAAAGFTFSSEYEMNVWRAPBKGLEWVISISSGSTIYY     | 60  |
| OY |  | 63  | ADTVKGRFTTIRSDNKNLTFLQMTSLRSEDVWYYCARDYGAVWGCGTTVTYSSGGGSGSg     | 122 |
| Dd |  | 61  | ADVXKGRFTTIRDNKNLSLYLQMSLRRAEDTAIVYICARPTTDYWGCGTTVTYSSGGGSGSg   | 120 |
| OY |  | 123 | GGSGGGGGSDIELTOSPAIMASAPGBERVMTTCASSSV-RYNMWFQOKSGTSPKRWIYDT     | 181 |
| Dd |  | 121 | GGSGGGGGSDIYMTCPSSTLSASVDGDRVITTCRASGGISSMILAWYOQKRGAPKVLIYKA    | 180 |
| OY |  | 182 | SKLSSGVAFRRFSGSGSGTSYSTLTISMEAEADAATYYCOOWSNPLTFEGAGTKLEIKR      | 238 |
| Dd |  | 181 | STLESQVSRFRFSGSGSGTGDTFLTLLTSLQPEDEPATIYYCOOYSYSTPWTFFGGGTYLEIKR | 237 |

```

Query Match      55.7%; Score 903.5; DB 15; Length 237;
Best Local Similarity 72.6%; Pred. No. 1e-53;
Matches 172; Conservative 27; Mismatches 37; Indels 1; Gaps 1;

QY      EVKLOESGGGGLVPGSGSRKLSCAASGFTFSFGHMWVRAPAEKGLIEWAYISGSSTIYY 62
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1 EVLGVGGGGGLVGQGSLRLSCAASGFTFSYEENWVRQAAPKGLEWYSISSGSTIYY 60

QY      63 ADYVKGFFITSRDNPKTLLFLOMTSLASEDTWMYICARDYGAVYGCGTTVTYSSGGGSG 122
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      61 ADVKGFFTISRDNANKSLYLQNMSLRAEDTAVYVCARDTDYVGQGLTVYSSGGGSG 120

QY      123 GGSGGGGGSPIDELTQSPAIMSAPGEKVMTWCASASSV-RYMNFPOOKSGTSPKRWYDT 181
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      121 GGSGGGGGSDIWTQSPSTISASVGDRTYTTCRASGGISMLANYQQKPKGAIPKLYIKA 180

QY      182 SKLSGVPAFPSSGSGTSYSTLTISMEADDAATYYCQOWSSNPLTFGAGTKLEIKR 238
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      181 STLESVGPFRSSGSGSGTDFTLTLTISIQPEDFATYYCQOYSTPTWTFGGTKLEIKR 237

RESULT 15
US-09-880-748-2005
; Sequence 2005, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OR INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PP523
; CURRENT APPLICATION NUMBER: US/09/880.748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2020
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2005

```





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 05:50:29 ; Search time 25 Seconds  
(without alignments) 943.564 Million cell updates/sec

Title: US-09-403-882a-2

Perfect score: 1622  
Sequence: 1 MAEVKLQESGGIVPGGSR.....VLLTIIILITLMQKKR 316

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents RA: \*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/1aa/PCUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description       |
|------------|--------|-------------|--------|----|-------------------|
| 1          | 1319   | 81.3        | 255    | 3  | US-09-553-498-8   |
| 2          | 1319   | 81.3        | 255    | 4  | US-09-518-869-8   |
| 3          | 1199   | 73.9        | 236    | 2  | US-08-190-199A-65 |
| 4          | 1185   | 73.1        | 240    | 2  | US-08-956-047-25  |
| 5          | 1102.5 | 68.0        | 223    | 2  | US-08-190-199A-63 |
| 6          | 964.5  | 59.5        | 284    | 3  | US-08-564-164A-2  |
| 7          | 943.5  | 58.2        | 225    | 2  | US-08-190-199A-61 |
| 8          | 919    | 56.7        | 553    | 3  | US-08-661-052-16  |
| 9          | 919    | 56.7        | 553    | 3  | US-09-188-082-16  |
| 10         | 919    | 56.7        | 553    | 3  | US-09-364-088-16  |
| 11         | 919    | 56.7        | 553    | 3  | US-09-102-716-16  |
| 12         | 907    | 55.9        | 353    | 4  | US-09-203-958A-4  |
| 13         | 900.5  | 55.5        | 281    | 3  | US-09-423-439-44  |
| 14         | 897.5  | 55.3        | 245    | 4  | US-09-138-091A-76 |
| 15         | 894    | 55.1        | 240    | 4  | US-09-192-854-2   |
| 16         | 890.5  | 54.9        | 245    | 3  | US-08-918-148-75  |
| 17         | 890.5  | 54.9        | 245    | 3  | US-08-918-148-78  |
| 18         | 890.5  | 54.9        | 245    | 4  | US-09-138-091A-73 |
| 19         | 881.5  | 54.3        | 245    | 4  | US-08-918-148-76  |
| 20         | 881.5  | 54.3        | 245    | 4  | US-09-138-091A-74 |
| 21         | 877    | 54.1        | 270    | 2  | US-08-652-507-2   |
| 22         | 876.5  | 54.0        | 238    | 4  | US-09-798-689-21  |
| 23         | 873.5  | 53.9        | 301    | 3  | US-08-661-052-14  |
| 24         | 873.5  | 53.9        | 301    | 3  | US-09-188-082-14  |
| 25         | 873.5  | 53.9        | 301    | 3  | US-09-364-088-14  |
| 26         | 873.5  | 53.9        | 301    | 3  | US-09-102-716-14  |
| 27         | 864    | 53.3        | 258    | 4  | US-09-526-738A-4  |

|    |       |      |     |   |                   |                    |
|----|-------|------|-----|---|-------------------|--------------------|
| 28 | 861   | 53.1 | 239 | 3 | US-08-279-772A-8  | Sequence 8, Appl1  |
| 29 | 861   | 53.1 | 239 | 3 | US-08-902-486-11  | Sequence 11, Appl1 |
| 30 | 861   | 53.1 | 282 | 2 | US-08-860-174A-10 | Sequence 10, Appl1 |
| 31 | 859   | 53.0 | 495 | 4 | US-09-948-004-18  | Sequence 18, Appl1 |
| 32 | 858.5 | 52.9 | 284 | 3 | US-09-184-658-40  | Sequence 40, Appl1 |
| 33 | 858.5 | 52.9 | 284 | 4 | US-09-504-262D-40 | Sequence 40, Appl1 |
| 34 | 855   | 52.7 | 244 | 4 | US-08-918-148-77  | Sequence 77, Appl1 |
| 35 | 855   | 52.7 | 256 | 4 | US-09-526-738A-2  | Sequence 75, Appl1 |
| 36 | 855   | 52.7 | 244 | 4 | US-08-918-148-75  | Sequence 75, Appl1 |
| 37 | 853.5 | 52.6 | 242 | 2 | US-08-553-497A-26 | Sequence 26, Appl1 |
| 38 | 852   | 52.5 | 599 | 1 | US-08-463-163-3   | Sequence 3, Appl1  |
| 39 | 851   | 52.5 | 246 | 1 | US-08-469-486-57  | Sequence 57, Appl1 |
| 40 | 851   | 52.5 | 246 | 2 | US-08-469-658-57  | Sequence 57, Appl1 |
| 41 | 850   | 52.4 | 666 | 3 | US-08-423-439-51  | Sequence 51, Appl1 |
| 42 | 847   | 52.2 | 222 | 2 | US-08-190-199A-67 | Sequence 67, Appl1 |
| 43 | 840.5 | 51.8 | 241 | 4 | US-09-581-345-5   | Sequence 5, Appl1  |
| 44 | 840.5 | 51.8 | 244 | 4 | US-08-553-497A-20 | Sequence 20, Appl1 |
| 45 | 838.5 | 51.7 | 242 | 2 | US-08-553-497A-28 | Sequence 28, Appl1 |

#### ALIGNMENTS

```
RESULT 1
US-09-553-498-8
; Sequence 8, Application US/09553498
; Patent No. 6309861
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: Process for the production of naturally folded and secreted protei
; FILE REFERENCE: Case 20379
; CURRENT APPLICATION NUMBER: US/09/553,498
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: EP99107412.1
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 8
; LENGTH: 255
; TYPE: PRT
; ORGANISM: E. coli
US-09-553-498-8

Query Match      81.3%; Score 1319; DB 3; Length 255;
Best Local Similarity 99.2%; Pred. No. 8.6e-91;
Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 MAEVKLQESGGIVPGGSRKLSGASGFTFSFGMHWROAPEKLEWVAYISSGSSSTI 60
DB      1 MAEVKLQESGGIVPGGSRKLSGASGFTFSFGMHWROAPEKLEWVAYISSGSSSTI 60

QY      61 YYADTYKGRFTTSRDNPKNKTLFLQMTSLRSBDTMYTCARDYCAVYGQGTTVVSSGGG 120
DB      61 YYADTYKGRFTTSRDNPKNKTLFLQMTSLRSBDTMYTCARDYCAVYGQGTTVVSSGGG 120

QY      121 SGGGSGGGGGSDIELTQSPALMSASPERVTMTCSASSVRVNMWFOQSGTSPKRWIYD 180
DB      121 SGGGSGGGGGSDIELTQSPALMSASPERVTMTCSASSVRVNMWFOQSGTSPKRWIYD 180

QY      181 TSKLSSGVAFARFSGSGGTSYSLTITSSMEADATYTCOOMSNPLTFAGTLELKRRA 240
DB      181 TSKLSSGVAFARFSGSGGTSYSLTITSSMEADATYTCOOMSNPLTFAGTLELKRRA 240

QY      241 AEQKLISEEDLNGA 254
DB      241 AEQKLISEEDLNGA 254

RESULT 2
US-09-618-869-8
; Sequence 8, Application US/09618869
```

Patent No. 6455279  
GENERAL INFORMATION:  
APPLICANT: Ambrosius, Dorthée  
APPLICANT: Rudolph, Rainer  
APPLICANT: Schaeffner, Joerg  
APPLICANT: Schwarz, Elisabeth  
TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND  
TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR  
TITLE OF INVENTION: CHAPERONES  
FILE REFERENCE: 20381  
CURRENT APPLICATION NUMBER: US/09/618,869  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: EP9911481.5  
PRIOR FILING DATE: 1999-07-23  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 8  
LENGTH: 255  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-618-869-8

Query Match 81.3%; Score 1319; DB 4; Length 255;  
Best Local Similarity 99.2%; Pred. No. 8, 6e-91;  
Matches 252; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAEVKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHVWROAPEKGLWVAIYSSGSSSTI 60  
DB 1 MAEVKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHVWROAPEKGLWVAIYSSGSSSTI 60  
QY 61 YVADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYCCARDYGAWGQGTTLTVSSGGGSG 120  
DB 61 YVADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYCCARDYGAWGQGTTLTVSSGGGSG 120  
QY 121 SGGGSGGGGSDIELTOSPAIMASPGERVMTCSASSSVRYNMWFOOKSGTSPKRWIYD 180  
DB 121 SGGGSGGGGSDIELTOSPAIMASPGERVMTCSASSSVRYNMWFOOKSGTSPKRWIYD 180  
QY 181 TSKLSSGVPARFSGSGSTSYSLTSSMEADATYYCQOMSSNPLTFGAGTKLEIKRA 240  
DB 181 TSKLSSGVPARFSGSGSTSYSLTSSMEADATYYCQOMSSNPLTFGAGTKLEIKRA 240  
QY 241 AEOKLISEEDLNGA 254  
DB 241 AEOKLISEEDLNGA 254

RESULT 3  
US-08-190-199A-65  
Sequence 65, Application US/08190199A  
Patent No. 5830663  
GENERAL INFORMATION:  
APPLICANT: EMBLETON, Michael J.  
APPLICANT: GOROCIOV, Guy  
APPLICANT: JONES, Peter T.  
APPLICANT: WINTER, Gregory P.  
TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS  
NUMBER OF SEQUENCES: 70  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILSBURY MADISON & SUTRO, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/190,199A  
FILING DATE: 13-JUL-1994

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/01483  
FILING DATE: 10-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9212419.7  
FILING DATE: 11-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9117352.6  
FILING DATE: 10-AUG-1991  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-190-199A-65

Query Match 73.9%; Score 1199; DB 2; Length 236;  
Best Local Similarity 96.2%; Pred. No. 6, 8e-82;  
Matches 227; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 EVKQESGGGLVOPGSGRKLSCAASGFTFSSFGMHVWROAPEKGLWVAIYSSGSSSTIY 62  
DB 1 DVQLVESGGGLVOPGSGRKLSCAASGFTFSSFGMHVWROAPEKGLWVAIYSSGSSSTIY 60  
QY 63 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYCCARDYGAWGQGTTLTVSSGGGSG 122  
DB 61 ADTVKGRFTISRDNPKNTLFLQMTSLRSEDTVMYCCARDYGAWGQGTTLTVSSAGGSG 120  
QY 123 GGGSGGGGSDIELTOSPAIMASPGERVMTCSASSSVRYNMWFOOKSGTSPKRWIYD 182  
DB 121 GGGSGGGGSDIELTOSPAIMASPGERVMTCSASSSVRYNMWFOOKSGTSPKRWIYD 180  
QY 183 KLSGVPARFSGSGSTSYSLTSSMEADATYYCQOMSSNPLTFGAGTKLEIKR 238  
DB 181 KLSGVPARFSGSGSTSYSLTSSMEADATYYCQOMSSNPLTFGAGTKLEIKR 236

RESULT 4  
US-08-956-047-25  
Sequence 25, Application US/08956047  
Patent No. 5882924  
GENERAL INFORMATION:  
APPLICANT: Fritz, Hans-Joachim  
APPLICANT: Hennecke, Frank  
APPLICANT: Kolmar, Harald  
TITLE OF INVENTION: Genetic Selection, by Means of Signal  
TITLE OF INVENTION: Transduction in Microorganisms, of Proteins which are  
TITLE OF INVENTION: Capable of Ligand Binding  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,047  
FILING DATE: 22-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/257,669  
FILING DATE: 08-JUN-1994  
APPLICATION NUMBER: DE P 43 19 296.3  
FILING DATE: 10-JUN-1993  
CLASSIFICATION: 435



US-08-564-164A-2

|                       |                  |                    |           |             |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match           | 59.5%;           | Score 964.5;       | DB 3;     | Length 284; |
| Best Local Similarity | 72.2%;           | Pred. No. 2.5e-64; |           |             |
| Matches 187;          | Conservative 24; | Mismatches 41;     | Indels 7; | Gaps 2      |

```

QY 1 MAEVKIQESGGGLVQCGSGRKLSCAASGTTFSFGMHVYVROAPEKGLFEMVYISSGSTI 60
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
26 MAQVYIQSGGGLVQGRSLKLSCVBSGTFPFTNYGMNMIKQTPKGLFEMVYISSGSYL 85
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 YYADYTKGRTISRDNPKNTLFLQMTSLRSBDTVMYCAR-----DYGAWYGGTTVY 114
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 86 YYAEYTKGRTISRDNKNTLYLQMTSLRSBDTALYYCARHGTGTGTFPDYWGCTTVY 145
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 115 SSGGGSGGGSGGGSGGSDIELTQSPALMASPDGEVMTWCASSGV--RYNMVFOOKGTS 173
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 146 SSGGGSGGGSGGGSGGSDVELTQSPHSLASLIGETVYIECLASBDISNYLAWYQOKGKS 205
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 174 PKRWYDTSKLSGVPARFSGSGSGSTYSYLTTSMEADDAITYYCOQWMSNPLTFGAGTK 233
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 206 POLLIYVASLQDGVSPRFSGSGSGTQFSLKISNMQPDGEVYVCOQAYKYPSTFGAGTK 265
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 234 LELKRAAEQKLISEEDLN 252
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 266 LELKRAAEQKLISEEDLN 284
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

## RESULT 7

US-08-190-199A-61  
; Sequence 61, Application US/08190199A

```

GENERAL INFORMATION:
APPLICANT: EMBLETON, Michael J.
APPLICANT: GOROCHOV, Guy
APPLICANT: JONES, Peter T.
APPLICANT: WINTER, Gregory P.
TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
City: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190.199A
FILING DATE: 13-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CB92/01483
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9212419.7
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9117352.6
FILING DATE: 10-AUG-1991
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
Type: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
IS-08-190-199A-61

```

|                       |        |                    |       |             |
|-----------------------|--------|--------------------|-------|-------------|
| Query Match           | 58.2%; | Score 943.5;       | DB 2; | Length 235; |
| Best Local Similarity | 75.0%; | Pred. No. 7.2e-63; |       |             |

Matches 177; Conservative 26; Mismatches 32; Indels 1; Gaps 1;

|    |   |     |
|----|---|-----|
| QY | 3 EKVLEDSGGGLVOPGSGRKLSCAASGFPPSSFGHNWRLOAPBKEGLEVAATIAISSCSSTIYY | 62  |
| Db | 1 QVLKESGPGFLWABPQSISTITCTVSGSRLTSYGHWNRORPPGKLEHLYIMAGGSF-NY     | 59  |
| QY | 63 ADTVKGRFTISHDNPKNTLPLQMTSLSESDTVMYCYCARDGYAWGCGTTVTVASGGGGSG   | 122 |
| Db | 60 NSALMSRLISIKDNSKSQVFLKNKNSLGTDPTAMTYCARDBGAYWGCGTLTVASGGGGSG   | 119 |
| QY | 123 GGGSGGGGSDIELTOSPAIMASPFGEKYTMTCASSSVRYMNMVFQQSGSGSPKRWIYDTIS | 182 |
| Db | 120 GGGSGGGGSOIVLTOSPAIMASPEQAKYTMTCASSSVSYTHMYHQXSGSGSPKRWIYDTIS | 179 |
| QY | 183 KLSGVPARFSGSGSGTSTYSLTISMEADAAITYCCOMSSNPLTFGAGTKLELR         | 238 |
| Db | 180 KLASGVPARFSGSGSATISYSLTISMEADAAITYCCOMSSNPLTFGAGTKLELR        | 235 |

## RESULT 8

US-08-661-052-16  
; Sequence 16, Application US/08661052

```

/ GENERAL INFORMATION:
/ APPLICANT: Vashwant M. Deo
/ APPLICANT: Joel Goldstein
/ APPLICANT: Robert Graziano
/ APPLICANT: Chazian Somaandaram
/ TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPTRISE
/ TITLE OF INVENTION: OF ANTI-PC RECEPTOR ANTIBODIES
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 60 State Street, Suite 510
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/661,052
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/484,172
/ FILING DATE: 07-JUNE-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Arnold, Beth E.
/ REGISTRATION NUMBER: 35,430
/ REFERENCE/DOCKET NUMBER: MXI-043CP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)227-5941
/ INFORMATION FOR SEQ. ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 553 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ JS-08-661-052-16

```

|                       |                  |                    |            |             |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match           | 56.7%;           | Score 919;         | DB 2;      | Length 553; |
| Best Local Similarity | 67.5%;           | Pred. No. 1.3e-60; |            |             |
| Matches 183;          | Conservative 22; | Mismatches 48;     | Indels 18; | Gaps 2;     |

Qy 2 AENVKQBSGGGIVNPQGSRKLSCAASGTFTSSFGMHVNRQAPENGLGVAVYISSGSTTY 61  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db 276 SPIKIQSGAGELVMRSCTSVKLSCTAASGNIKDSIMHILRLRGPGEGLGMIGHIDPENODTE 335

Qy 62 YADTVYKGRFTISRDPNKKITLFLQMITSLSBEDTVVVYYCARD-----YGAYMGCGTITYVS 115

Db 336 YAPKFGKATFTTDTSSNTAYVLQSLTSEDTAVVYCNCTPTGPPYFDYWGQTTVTS 395  
Qy 116 SGGGGGGGGGGGGGGGDIETLTPSPAIMSASPGERVMTTCASSSVRYMMFOOKSGTSPK 175  
Db 396 SGGGGGGGGGGGGGSENVLTQSPAIMSASPGKVTITCASSSVRYMMFOOKPGTSPK 455  
Qy 176 RWIYDTSKLSGVPAFPSSGSGTSTSLTSSMEADATYYCOQMSNPFTGAGTKLE 235  
Db 456 LWTYSTNLASGVPAFPSSGSGTSTSLTSSMEADATYYCOQMSNPFTGAGTKLE 515  
Qy 236 LKR-----AAEQKLISEEDLNGA 254  
Db 516 LKRAAGSGGGGSDIAAEQKLISEEDLNGA 546

## RESULT 9

US-09-188-082-16  
Sequence 16, Application US/09188082  
Patent No. 6270765  
GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo  
APPLICANT: Joel Goldstein  
APPLICANT: Robert Graziano  
APPLICANT: Chezian Somasundaram  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09188, 082  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/661, 052  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MXI-043CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-188-082-16

Query Match 56.7%; Score 919; DB 3; Length 553;

Best Local Similarity 67.5%; Pred. No. 1,3e-60;  
Matches 183; Conservative 22; Mismatches 48; Indels 18; Gaps 2;

Qy 2 AEVYKQSGGGLVQPGSRKLSGASGFTFSFGMHVWQAPKGLWVAAYISSGSGTIIY 61  
Db 276 SDIKLQSGAEIVRSSTVSLCTASGFNIKOSYMHMLQPGLEWIGMDPENGDTE 335  
Qy 62 YADTVAGRTISNDKNTLFIQMTSLRSEDVTMYVCARD-----YGAYWGQTTVTS 115  
Db 336 YAPKFGKATFTTDTSSNTAYVLQSLTSEDTAVVYCNCTPTGPPYFDYWGQTTVTS 395

Qy 116 SGGGGGGGGGGGGGGGDIETLTPSPAIMSASPGERVMTTCASSSVRYMMFOOKSGTSPK 175  
Db 396 SGGGGGGGGGGGGGSENVLTQSPAIMSASPGKVTITCASSSVRYMMFOOKPGTSPK 455  
Qy 176 RWIYDTSKLSGVPAFPSSGSGTSTSLTSSMEADATYYCOQMSNPFTGAGTKLE 235  
Db 456 LWTYSTNLASGVPAFPSSGSGTSTSLTSSMEADATYYCOQMSNPFTGAGTKLE 515  
Qy 236 LKR-----AAEQKLISEEDLNGA 254  
Db 516 LKRAAGSGGGGSDIAAEQKLISEEDLNGA 546

## RESULT 10

US-09-364-088-16  
Sequence 16, Application US/09364088  
Patent No. 6365161  
GENERAL INFORMATION:  
APPLICANT: Yashwant M. Deo, et al.  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED  
TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street, 24th Floor  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09364, 088  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/188, 082  
FILING DATE: 07-JUNE-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/484, 172  
FILING DATE: 07-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remillard, Jane E.  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: MXI-043CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)742-7414  
TELEFAX: (617)742-7414  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-364-088-16

Query Match 56.7%; Score 919; DB 3; Length 553;

Best Local Similarity 67.5%; Pred. No. 1,3e-60;  
Matches 183; Conservative 22; Mismatches 48; Indels 18; Gaps 2;

Qy 2 AEVYKQSGGGLVQPGSRKLSGASGFTFSFGMHVWQAPKGLWVAAYISSGSGTIIY 61  
Db 276 SDIKLQSGAEIVRSSTVSLCTASGFNIKOSYMHMLQPGLEWIGMDPENGDTE 335  
Qy 62 YADTVAGRTISNDKNTLFIQMTSLRSEDVTMYVCARD-----YGAYWGQTTVTS 115  
Db 336 YAPKFGKATFTTDTSSNTAYVLQSLTSEDTAVVYCNCTPTGPPYFDYWGQTTVTS 395  
Qy 116 SGGGGGGGGGGGGGGGDIETLTPSPAIMSASPGERVMTTCASSSVRYMMFOOKSGTSPK 175  
Db 396 SGGGGGGGGGGGGGSENVLTQSPAIMSASPGKVTITCASSSVRYMMFOOKPGTSPK 455







Qy 179 YDTSKLSGVPAAPFGSGSGSTSYSLTISMEADATYYCOOMSSNP LTFGAGTKLELKR 238  
Db 181 YAASSLQSGVPSRFSGSGSGTDFTLTISLQPEDFATYYCOO SYSTPNTFGQGTKEIKR 240

Search completed: June 23, 2005, 05:55:04  
Job time : 67 secs